
Measuring the Burden of Work Disability: A Review of Methods, Measurement Issues and Evidence

4

Emile Tompa

This chapter describes issues and measures related to the burden of work disability, including both direct costs (i.e. health care, wage replacement benefits and rehabilitation services from various public and private insurance providers) and indirect costs (i.e. labour productivity and output losses).

4.1 Introduction

This chapter describes issues and measures related to the burden of work disability in a variety of developed countries. The burden encompasses working age adults whose engagement in the labour force is temporarily or permanently compromised due to a health condition. The burden includes both direct costs (i.e. health care, wage replacement benefits and rehabilitation services from various public and private insurance providers) and indirect costs (i.e. labour productivity and output losses). The full extent of the burden of disability can and does encompass nonfinancial and sometimes intangible outcomes such as indi-

vidual role functioning outside of the paid labour force and health-related quality of life. We touch on these matters, but focus on work disability burdens.

A number of studies have investigated the financial impact of specific health conditions such as low back pain, depression, arthritis and diabetes, but fewer studies have considered the financial impact of all health conditions. The few that have, find the costs to be substantial. For example, in the United States (US), health-related lost productivity was estimated at \$226 billion/year or \$1,685/employee per year in 2002 (Stewart et al. 2003). The largest proportion of this cost, 71%, is attributable to reduced performance while at work (i.e. presenteeism). These estimates consider only a fraction of the cost of work disability because they only include individuals actively engaged in paid employment. Not included is lost output associated with individuals who were not employed due to a work disability. Estimates have been made for Canada that attempt to capture both short-term and long-term disability, the latter which includes individuals not actively engaged in the labour force (Health Canada 1989, 1996, 1998). For 1998 the estimate is \$16.9 billion or 1.05% of gross domestic product (GDP) (Health Canada 1998). This amount does not include the costs of presenteeism as does the US estimate. Neither the US nor Canadian estimates include the value of lost productivity in social roles outside of work, nor the health care and other related costs associated with the various conditions that gave rise to the disabilities.

E. Tompa (✉)
Scientist/Health and Labour Economist,
Institute for Work & Health, 481 University Avenue,
Suite 800, Toronto, ON, Canada M5G 2E9

Department of Economics, McMaster University,
Hamilton, ON, Canada

Dalla Lana School of Public Health,
University of Toronto, Toronto, ON, Canada
e-mail: ETompa@iwh.on.ca

Another way to assess the burden of work disability is through the number of working aged individuals who are not employed/out of the labour force due to a work disability, or the converse, how many are employed. Canadian estimates for the time period 1989–2001 of the proportion of people with disabilities employed at some point during the year are much smaller than for people without disabilities, 43–63% compared to 84–88% (Tompa et al. 2006). Fewer still are in the labour force all year and/or employed all year, 43–58% for the former, and 29–49% for the latter. These trends have remained relatively stable over the time period considered in that study.

Chronic conditions contribute enormously to the work disability burden, and are likely to increase in proportion as the population ages in many developed countries. A study in the USA for 1995–1996 found that 6.7 days per month were lost by individuals with impairments aged 25–54 (Kessler et al. 2001). This aggregates to 2.5 billion work-impairment days per year. The major conditions found to be contributing to impairment days in the USA were cancer, ulcers, major depression and panic disorder.

From the above examples, it is clear that the burden of work disability can be substantial in developed countries. Undoubtedly, the issue of work disability and its prevention warrants attention by governments at all levels and society at large, since there is much to be gained in terms of productivity and growth opportunities if the burdens can be appropriately addressed.

In this chapter we provide an overview of the burden of work disability in financial terms and with other measures. We begin with an explanation of the value of measuring the burden of disability and specifically work disability. This is followed with a brief discussion of the importance of evaluating the economic returns of work disability prevention initiatives. This topic is elaborated upon in Chap. 23 of the handbook. We follow with an overview of disability prevalence and benefits receipt across OECD countries and a description of the extent of the burden for several developed countries. We conclude with a summary and suggestions for the way forward.

4.2 Burden Studies and Their Role

Burden of disease studies measure the total value of lost healthy time (i.e. morbidity and mortality) from a particular disease or health condition, the costs of treating individuals with the condition, the cost of other services provided due to illness and disability and the impact of the health condition in terms of lost output and productivity. Though burden studies do not measure the probability of success of treatment options or the opportunity costs of interventions that might be undertaken to reduce the burden, these types of studies serve an important information role. They provide insights into the magnitudes of the health loss and the cost of a health condition to society. This information can be used to assess how burdens may have changed over time, how they compare to burdens for similar conditions in other jurisdictions or how they compare to the burden of other health conditions. Such information can be invaluable to policymakers for priority-setting purposes. Burdens that appear particularly onerous may bring attention to the need for (1) increasing funding for intervention options known to reduce the burden, (2) evaluating the merits (both in terms of health and resource implications) of burden reduction resulting from known treatment options that have not yet been evaluated and (3) investing in research to discover treatment options to reduce the burden in cases where no new alternatives currently exist.

Estimates of expenditure (e.g. health care costs) in burden studies are typically assessed for a specific calendar year and are based on costs in that year for all individuals diagnosed with or living with a particular health condition. These aggregate costs are also referred to as prevalence costs, because they encompass costs for individuals across the health trajectory, including the newly diagnosed, long-term survivors as well as those at the end of life. Burden studies can also report health care costs longitudinally, starting from diagnosis, and only include newly diagnosed patients. The time period for these longitudinal or incidence cost studies ranges from several months to the patient lifetime following diagnosis. These two general

types of burden studies are not directly comparable, because of differences in the time periods measured and the inclusion criteria.

One method of modelling health care costs for a particular health condition is the phase of care approach. This approach divides services and costs following diagnosis into distinct periods or phases (e.g. initial, continuing, last year of life) and can be used to estimate either incidence or prevalence costs. When phase of care-specific cost estimates are applied to survival probabilities for an incident cohort, the result is analogous to an incidence cost estimate. When phase of care-specific cost estimates are applied to phase-specific person-years of survival within a specific year, the result is a prevalence cost estimate.

Costs incurred by disability insurance providers for wage replacement are generally not included in societal/country-level burden studies because such compensation is considered a transfer of purchasing power from one group of individuals to another, rather than an expenditure of resources by society. Nonetheless, they are of relevance at the disability system level, and the magnitude of these costs is an important consideration for insurance providers, whether public or private. Additionally, insurer costs associated with service provision, such as health care, return to work coordination and physical and vocational rehabilitation, need to be included in the burden estimate, since these are truly expenditures of resources associated with the treatment of a health condition.

Estimates of lost output and productivity (sometime labelled ‘indirect costs’) associated with work disability from a health condition are also assessed for a specific calendar year when using the prevalence approach. For this approach, estimates are based on output/productivity losses in that year for all individuals diagnosed with or living with a particular health condition. How these estimates are calculated is less standardized than the measurement of other expenditures. However, two approaches are commonly used—the human capital approach and the friction cost approach. These two approaches are elaborated upon later in this chapter (See also Chaps. 3 and 23). To estimate the burden of premature mortality from a health condition, the output or

indirect costs over the remainder of the forgone working career are generally included in the estimate. Here too the human capital approach or the friction cost approach can be used. The two approaches diverge substantially in their estimates of indirect costs for such incidents.

Estimates of lost output and productivity using the incidence approach require calculating the lifetime losses associated with all new cases of a health condition. Here too the human capital approach or the friction cost approach can be used. As noted, the two approaches will have dramatically different estimates, particularly if the health condition being evaluated has long-run disability implications.

4.3 Concepts and Measurement of Disability

Disability, and specifically work disability, is associated with a health condition, but is more than just the existence of a condition itself. There is an extensive theoretical literature on disablement that is largely centred on two conceptual frameworks: those of Nagi and the World Health Organization (WHO). Nagi’s work from the 1960s is one of the first comprehensive conceptualizations of disablement (Nagi 1965). In his framework, disablement is a series of four interrelated concepts that describe the impact of a health condition on a person’s body, activities and involvement in society (Nagi 1965, 1991). These four concepts are pathology, impairment, functional limitation and disability. Disability can arise from a functional limitation or directly from an impairment. But not all impairments and functional limitations give rise to disability. A critical factor is the degree to which the social environment creates barriers to involvement for an individual with an impairment or functional limitation.

The WHO developed a conceptual framework for disablement comparable to, but independent of, the Nagi model (World Health Organization 1980, 2001). This framework describes the consequences of disease as four interrelated concepts: disease (health condition), impairment (body structure/function), disability (activity) and handicap (participation). The WHO model

appears similar to that of Nagi, with the disease (health condition) dimension comparable to Nagi's pathology, the term impairment (body) being used in both models for the second concept, the disability (activities) dimension comparable to functional limitations and handicap (participation) to disability. Both frameworks are based on the notion that disability is not a characteristic of an individual, but a relational concept that derives from the interaction of an individual's abilities and other personal characteristics with a particular social and built environment.

The socio-medical concept of disability described above is difficult to operationalize, particularly for the purpose of assessing partial work disability. Consequently, many disability compensation programmes only compensate for total disability. One of the few exceptions is workers' compensation programmes which offer fractional pensions based on various formulas for assessing partial disability (Dembe 2000; Peterson et al. 1998). In general, workers' compensation programmes use one of four approaches to compensate for permanent impairment: (1) a medical assessment of the degree of permanent impairment, (2) an estimate of loss of wage-earning capacity, (3) an estimate of actual wage loss or (4) a hybrid of the former three (Pauly et al. 2002). Three types of hybrids are common: (1) one that uses a different system for different types of injuries, (2) one that pays both impairment benefits and benefits for loss of wage-earning capacity or actual wage loss or (3) a system in which the same injury can lead to either an impairment-based benefit or a benefit based on loss-of-wage-earning capacity or actual wage loss.

The measurement of burdens from health conditions and related disability generally focuses on financial metrics. But burdens can also be depicted with nonfinancial data such as the number of cases in a population, the severity of cases, and for work disability, the number of individuals absent from work/unemployed, out of the labour force or receiving disability benefits. Statistics on the number of people with disability is often assessed through self-reported health survey. Employment, unemployment and out of the labour-force statistics are often drawn from self-reported labour-force surveys, which are undertaken on a monthly

basis in many developed countries. Some such surveys also inquire about reasons for unemployment or disengagement from the labour force, with one of the categories being injury, illness or disability. Census data may also provide the relevant information, but is less frequently collected. Counts of the number or proportion of individuals who are disability benefits recipients are generally developed from administrative data sources associated with the various disability compensation programmes. Such data cannot provide an accurate estimate of the number or proportion of disability individuals in a population, since programmes may not be universal. Furthermore, not all eligible individuals may apply or receive benefits for a variety of reasons. Interpreting differences in disability reciprocity across countries is particularly a challenge, because the criteria for eligibility may vary dramatically from country to country. The World Bank has developed a metric exclusively designed to estimate the burden of disease in society, known as the disability-adjusted life year (DALY) (Murray 1994). See Sect. 4.3.1 below for details on this construct.

4.3.1 Disability-Adjusted Life Years

A DALY is a time-based measure of the burden of disease that combines years of life lost from premature mortality and years of life lived in less than perfect health. Age weights are used for the value of time at different ages to reflect the dependence of the young and older individuals on working age adults. These weights are associated with societal values of productivity and investment in education. Severity weights for different disability states are also used to adjust the value of time with health conditions. These weights are between zero and one and are based on a value of death as zero and perfect health as one. Future DALYs are discounted to the present time using a discount rate, customarily 3% (World Health Organization 2011).

The DALY measure can provide a comparable metric for assessing burdens across different categories of health conditions, or across different countries. For example, Polinder et al. (2007) uses DALYs to compare injury-related burdens across six European countries. Because DALYs are constructed in a unique way with underlying assumption built into their weighting system, they are not readily comparable to monetary measures of burdens. In particular, indirect costs (generally associated with productivity losses) are accounted for through standardized age weights with DALYs, rather than based on actual measurement of losses.

4.4 Measuring Societal-Level Indirect Costs Associated with Work Disability

Some indirect costs of adverse health and related work disability can be immediate (e.g. lost output due to sickness absence), while others unfold over longer periods of time (e.g. reduced capital accumulation due to reduced savings over the life course). One of the principal indirect costs associated with adverse health of the working age population is reduced productivity and output. The effect of health on labour-force participation and earnings is sometimes described as health as a capital or an investment good, because it is seen as a stock of capital that one can draw on over time to earn a livelihood (Grossman 1972). Reduced productivity and output associated with health may arise through health-related absenteeism and presenteeism, or reduced labour-force engagement such as unemployment or nonparticipation due to poor health (Sharpe and Murray 2010). More generally, health may affect labour quality, i.e. healthy adults have higher energy levels and mental acuity than less healthy adults, and therefore may be more productive. At the organizational level, absenteeism and presenteeism may affect team productivity and output (Nicholson et al. 2006; Pauly et al. 2002). Other contributions at the organizational level to output, such as social contribution (i.e. payroll taxes) and profits, may also be affected by lower levels of productivity and output as measured by the wages of workers.

Longer run pathways by which adverse health and disability may affect productivity and output include child health and its association with educational attainment, reduced saving and its implications for capital accumulation and socio-demographic factors such as fertility levels and female participation in the paid labour force (Bloom and Canning 2000; Bloom and Sachs 1998). Premature mortality will also affect labour-force size and output. Sharpe and Murray (2010) suggest that for developed countries, only the first of these longer run pathways is likely to be relevant. Table 4.1 summarizes the various pathways by which health and disability might impact output.

Poor health can also compromise participation in activities outside of paid work. These roles may include parenting, home maintenance, community involvement, religious activities and leisure activities. The impact of health on such participation might be described as health as a consumption good, as per Grossman (1972). The Grossman model of the demand for health, which is used widely in health economics, is less refined about social roles outside of the paid labour force, since it is designed around the traditional economic paradigm of work and leisure. A more holistic approach to the impact of health on individuals was described above, i.e. the Nagi (1965, 1991) and the WHO (1980, 2001) frameworks. Good health also has intrinsic value in and of itself. Being healthy allows one to enjoy life more fully in all social roles, whether in the paid labour force or outside of it. This intrinsic value of good health is sometimes called health-related quality of life, and would also be put under the category of health as a consumption good.

Time spent seeking care may also take individuals away from paid work and/or participation in other social roles. Further, other individuals in the family unit and in the community may be affected by an individual's health. Family, friends and neighbours may provide informal caregiving. There may also be some substitution in the roles of family members, such as a spouse entering the paid labour force if an individual is unable to participate in this role due to poor health. Quantifying the monetary value of time spent seeking care and time use of other individuals can be a challenge.

Table 4.1 Summary of pathways from health and disability to output via the paid labour force

Adult health and output	Current health → presenteeism, absenteeism, employment, labour-force participation, size of the labour force – Output per hour due to presenteeism (team production may also be affected) – Output per person due to absenteeism (team production may also be affected) – Output per labour-force participant due to health-related unemployment – Output per working age population due to health-related nonparticipation – Size of the labour force due to premature mortality
Child health, educational investment and output	Child health → educational attainment → human capital → productivity and output over the life course
Life expectancy, savings and capital investment	Life expectancy → savings for retirement → capital investment → productivity and output
Child health and demographic effects	Child health → fertility → size of the working age population → output Child health → fertility → female participation in paid labour force → output

Table 4.2 Aspects of indirect costs of health and disability

Output of paid labour force	Adult health, productivity and output (including organizational and societal-level effects) Child health, educational attainment, productivity and output Savings, productivity and output Demographics, fertility, mortality, size of the paid labour force and output
Participation in roles outside of paid work	Parenting Home care Community involvement Religious activities Leisure activities Education
Health-related quality of life	Intrinsic value of good health
Time use of other individuals	Family/community time in caregiving Family role substitution

To summarize, Table 4.2 below highlights the various aspects of indirect costs of adverse health and disability.

Estimating the total burden of adverse health and disability across all the above-noted categories is a substantial measurement task. Consequently, many burden studies focus only on the indirect costs associated with loss of output/productivity of adults experiencing the condition, as well as the loss of health-related quality of life of all individuals with the condition. Time use of other individuals in the family and com-

munity would also be relevant in such cases, but would likely be of a smaller magnitude and are less often included.

4.5 Measuring Indirect Costs of Adult Onset of Disability

Figure 4.1 below depicts how one might classify productivity and output losses from injury or illness at the individual level. The schema distinguishes between temporary impairment, permanent

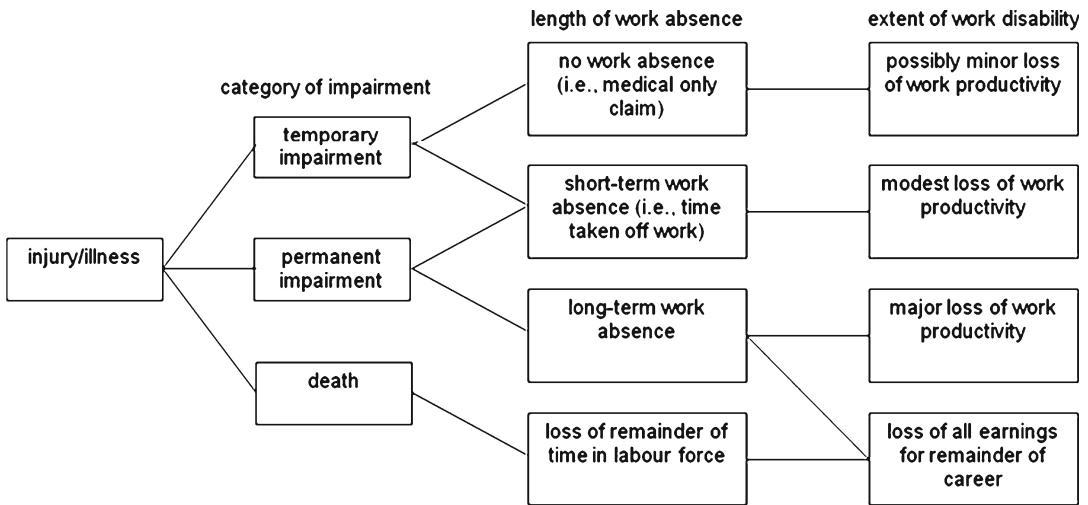


Fig. 4.1 Schema for categorizing work disability (adapted from Weil 2001)

impairment and fatality, as well as between absences and loss of work output and productivity at the individual level. Not depicted in the schema is the possibility of long-term losses through non-disability labour-market outcomes such as job loss or loss of promotion that can lead to further losses over one's career. This may be associated with any category of impairment, work absence or work disability. Many individual losses are not identified in the schema. Impairment from a work injury or illness may result in nonwork disability (i.e. disability in other social roles). Nonwork losses associated with impairment from a work injury or illness may also include loss of health-related quality of life (intrinsic value of health). There may also be uncompensated out of pocket costs associated with health loss.

A key controversy in measuring the burden of injury/illness and disability is how to measure the value of lost output and productivity associated with long-term work disability. Historically, the human capital approach was used to measure productivity losses associated with work disability. The approach assumes full employment (usually only implicitly), and that it is impossible to replace injured or ill workers from the ranks of the unemployed. In the absence of an intervention, productivity losses are assumed to continue until return to work, or in the case of permanent work disability and death, until age of retirement. The approach

has a very strong zero-substitution assumption that is more or less permanent. Koopmanschap et al. (1995) describe the human capital approach as a measure of potential productivity losses. It might best be considered an upper-bound estimate of the long-term burden. In the short term, losses might actually exceed the wage cost of the absence due to the disruption in the production process resulting from the occupational injury or disease.

Essentially, the human capital approach is an estimate of the counterfactual, that is, what the individual would have earned or produced had they not been injured or ill. Actual wages are used to calculate labour-market losses and assumed to be either fixed over time (this is the basis on which many workers' compensation wage replacement programmes operate) or adjusted for lifetime earnings growth. Adjustments are based on data from population statistics (stratified where desired by occupation, educational attainment and other relevant labour-market earnings characteristics) or collected through matching of injured individuals with a healthy cohort on socio-demographic characteristics and contextual factors that bear on earnings potential (see Weil 2001 for a summary of methods). For nonwage work, the opportunity cost of time or replacement cost approach might be used to estimate potential productivity losses in these roles (see Drummond et al. 2005, p. 216 for details).

In considering how a work absence might affect a firm's productivity, we describe the key factors that might bear on the magnitude of these consequences. A firm's adjustment to an occupational injury or disease can be achieved in various ways, depending on the nature of the production process and the duration of absence (short-term vs. long-term). With short-term absences, some work can be postponed, some might be taken over by colleagues (during regular work hours or on an overtime basis) and some might be completed by a replacement worker from internal labour reserves or from a temporary employment agency. With longer-term absences, a temporary or permanent replacement may be hired or the extra work can be distributed among the existing staff by cutting less time-sensitive work.

If the firm maintains its production rates during the early period of the absence, it may incur additional costs such as overtime payment for other employees, a premium for temporary replacement workers or the costs of hiring a permanent replacement worker and associated incremental costs such as training costs. The total value of these productivity related consequences during this period will consist of the value of lost production (if any), the additional labour costs and recruiting and training costs. The length of the early period of losses will depend on the state and efficiency of the labour market as well as the occupation of the injured worker, the industry in question and the associated learning time required for a new recruit to get up to speed. If the level of unemployment in the economy as a whole is higher than the level of frictional unemployment, firms will be able to replace injured workers more readily. Identifying whether the unemployment rate exceeds the frictional level may be a challenge; though on average unemployment has been sufficiently high in most developed economies over the last two decades for it to have probably exceeded the frictional rate for much of the time. Additionally, firms have increasingly relied on flexible hiring practices, such as temporary and on-call contracts and temporary employment agency hires, thus providing them with a pool of backup labour to adjust to market shocks (see Tompa et al. 2007 for review of the literature).

The 'friction cost approach' (Koopmanschap et al. 1995) is one approach to measuring the productivity consequences of health improvements at the aggregate level. This approach is discussed in more detail in Chap. 3. According to this approach there is a short-run friction period during which a firm and society may incur losses as an adjustment is made to a worker's absence. In the long run no losses are held to occur because the injured worker either returns to work and performance returns to the pre-injury level or the firm replaces the injured worker with a new hire and performance eventually becomes comparable to what it was before. Table 4.3 provides an example of an analysis using both the friction cost and human capital approaches applied to data from the Netherlands. As is apparent, the friction cost approach consistently identifies much smaller productivity losses than the human capital approach.

Friction costs methods are likely most appropriate for marginal changes in absenteeism and work disability associated with health and safety interventions. When considering substantial changes in labour-force participation, such as an increase in the engagement of working age individuals with disabilities, the impact on the macroeconomic environment can be substantial and would require a macroeconomic model to estimate the effect on the general equilibrium of an economy. Most burden studies do not take this approach because of the computational challenge associated with estimating a general equilibrium model. Rather, they rely on the human capital approach under the implicit assumption that it is an acceptable first-level approximation based on the vantage point of the existing situation.

More recent work on the productivity costs of health has focused on including consideration of losses from presenteeism (Brouwer et al. 2002). A variation of the friction cost approach recognizes that a period of reduced performance might occur before and/or after a health-related work absence, or that there might simply be a period of reduced performance without an absence. Brouwer et al. (2002) examined data from a Dutch trade firm and found that productivity losses due to reduced performance at work accounted for about

Table 4.3 Example of the divergence between friction costs and human capital approaches

Cost category	Friction cost approach	Human capital approach
Absence from work	9.2	23.8
Disability	0.15	49.1
Mortality	0.15	8.0

Koopmanschap et al. (1995) used data from the Netherlands and compared the friction cost and human capital approaches to demonstrate the divergence in values derived under each approach. A comparison is made of the indirect cost of disease in the Netherlands in 1988 in billions of Dutch guilders.

14% of total productivity losses. The literature on measuring health-related at-work performance (presenteeism) is still relatively young, though its volume is growing rapidly.

4.6 Economic Returns of Work Disability Prevention Initiatives

As noted, burden studies provide policymakers with a sense of the magnitude of losses associated with a particular health condition, and provide insight into what might be gained if interventions available to address it are implemented. If a particular burden is deemed sufficiently large to warrant attention, the next step may be a search for promising interventions to reduce it. In some cases two or more alternatives may be considered. Alternatives under consideration ought to be evaluated for both their effectiveness and cost-effectiveness before they are adopted across the board. The economic evaluation of alternatives is an important part of programme evaluations. In this handbook, Chap. 23 focuses on providing guidance on the methodology of economic evaluation with a particular focus on its application to work disability prevention programmes. There is a growing literature on this topic. A recent systematic review of the literature found strong evidence in support of such intervention based on their financial merits (Tompa et al. 2008). Most of the better quality studies identified in the review took a system- or societal-level perspective, were coordinated through an insurance provider or workers' compensation authority and served a multi-sector client base. The programmes were multifaceted, offering a range of services such as return to work

coordination, ergonomics worksite visits, physiotherapy, behavioural therapy, rehabilitation and educations. Chapters 20 through 22 of this handbook provide details on a range of clinical, workplace and complex interventions designed to reduce work disability.

4.7 Disability Prevalence and Benefits Receipt Across OECD Countries

Statistics on the prevalence of disability in OECD countries suggest that it is a relatively common phenomenon. On average, approximately 14% of individuals report a chronic health condition or a disability across OECD countries (OECD 2010). The percentage varies from country to country, ranging from upwards of 20% in Estonia to just over 5% in Korea. A focus on work disability rather than disability associated with any social role will likely reduce percentages, since some individuals with health conditions may be employed in the labour market. Prevalence information provides a first-level approximation of the burden of disability across countries, but comparability is an issue because surveys used to estimate these statistics in different countries use different questions to inquire about health and function. Differences in cultural norms and other contextual factors may also influence perceptions and reporting even if similar questions are used.

One approach to estimating the prevalence of work disability is to identify the unemployment rates of people with disability. Generally, unemployment rates in this group are twice as high as for able-bodied individuals—14% on average in OECD countries compared to 7% for the non-dis-

abled (OECD 2010). Unemployment rates do not include individuals who have given up seeking work or who have exited the labour force entirely. To address this concern, another approach to estimating the prevalence of work disability is to compare the employment rates of disabled people as a percentage of all disabled working age adults compared with their able-bodied counterparts. In general, employment rates of people with disability are lower than for people without disabilities. Across 27 OECD countries employment rates for the disabled averaged approximately 44% compared to 75% for people without disabilities (data is for late 2000s, i.e. just prior to downturn in the global economy) (OECD 2010). What is not captured in these numbers is the level and type of engagement in paid work. Some employed individuals may be under employed, both in terms of hours worked and in the match between skill level and job challenges. In fact, the disabled are significantly more likely to be working part time than non-disabled employed individuals.

Low employment rates of people with disabilities are particularly a concern, given the aging of the population. For example, in some countries such as Italy, Japan, Korea and Spain, more than 1/3 of the population is projected to be over age 65 by 2050 (OECD 2010). Projections for other developed countries are also high. For example, Canada is projected to have 1/4 of the population over age 65 by 2050 (Human Resources and Skills Development Canada (HRSDC) 2011) and the USA 1/5 (Department of Health and Human Services (DHHS) 2011).

As might be expected, the disabled have lower incomes—between 15 and 30% lower (OECD 2010). Incomes are particularly lower than their able-bodied counterparts in English-speaking countries, whereas the differences are less remarkable for Nordic countries (less than 10%). The disabled also have a significantly higher probability of poverty, 22% compared to 14% for people without disabilities (OECD 2010). Poverty levels amongst the disabled are particularly high for the USA, Australia, Ireland, Korea and Canada. There is little difference in the risk of poverty in Sweden, Norway, the Netherlands and Slovakia.

A measure associated with the burden of work disability is the number of individuals receiving disability benefits. In 2007, the overall disability reciprocity rate in OECD countries was 6%, with higher rates in Hungary, Norway and Sweden (approximately 10%), and low rates in the non-English-speaking OECD countries of Japan, Korea and Mexico (below 2%) (OECD 2010). In general, countries with more universal programmes had higher rates. For example, northern European countries had rates between 8 and 11%, whereas the Anglo-Saxon countries, where eligibility is more limited, had rates in the 5–7% range. Disability benefit reciprocity rates are generally much higher for older workers, and even more so in countries where it serves as a transition to retirement.¹ On average, more than half of disability benefits recipients are men, though in Nordic countries the majority is women (OECD 2010). It is important to note that reciprocity rates may vary from country to country for reasons other than the prevalence of disability. In particular, the types of programmes provided and their eligibility rules can vary dramatically.

Disability benefit reciprocity rates have been increasing in many OECD countries over the last three decades, but are relatively stable in most recent times (OECD 2010). In particular, some countries have introduced policy changes in an effort to reduce disability inflows (e.g. Poland, Portugal, Luxembourg and the Netherlands). These policy changes in OECD countries are discussed in detail in Chap. 22. Such efforts, unless accompanied by labour-market reintegration programmes, may exacerbate unemployment and poverty rates for individuals with disabilities. Particularly noteworthy is the fact that the inflow into disability benefits on the grounds of mental health conditions has been rising in many OECD countries, and has become the leading cause in

¹In the Netherlands disability benefits reciprocity was quite high in the 1990 before the introduction of reforms to reduce the use of the programme as a substitute for unemployment or a transition to retirement. The Dutch experience with these reforms is described in de Jong and de Vos (2005) and de Vos et al. (2012).

many countries. The proportion of younger recipients has also been rising.

With the aging of the populations in many countries, reciprocity rates are likely to continue to increase, all else being equal. This is particularly apparent if one looks at cross-sectional statistics on the number of people on disability benefits at older ages. For men, the proportion on disability benefits more than doubles (and in some cases triples) between the ages of 45 and 64 (Milligan and Wise 2012). For example in Italy, Spain, Germany, Sweden, Belgium, Canada, France and the Netherlands, it is under 5% for men aged 45 but increases at age 64 to over 35% for Sweden, over 25% for the Netherlands and over 20% for Germany.

Generally, few people leave disability benefits programmes; benefits serve as a permanent source of income replacement. Only around 1–2% of recipients leave for reasons other than death. The largest outflows are in the UK, New Zealand and Australia, where over 5% of beneficiaries left reciprocity status for reasons other than death in 2008 (OECD 2010). Oddly, only a small fraction of outflow is into employment, specifically between 10 and 20% of total outflow.

As noted, data on reciprocity rates fails to account for the fact that many disabled individuals do not receive disability benefits. In fact, only a minority receive benefits. On average it is 25%, with the proportion as low as 10–15% in Portugal and Germany and as high as 33% for Norway, Poland and the USA (OECD 2010). Higher rates do not necessarily imply higher incomes, since generosity of benefits varies from country to country. Furthermore, some disabled individuals may also receive other types of benefits, such as unemployment insurance. The proportion not receiving any benefits is 10–25% on average, but as high as 50% for some English-speaking and Mediterranean countries (specifically Canada, the USA, Spain, Greece) (OECD 2010). Some of these disabled individuals not receiving any benefits may be employed. Between 10 and 20% from these four countries have no public pension or labour-market income. For most OECD countries the proportion of no pension or labour-market income is less than 10%.

4.8 Financial Burden of Work Disability

The average spending of public disability programmes (including public sickness benefits) for OECD countries was 1.2% of GDP in 2007 (OECD 2010). Disability and sickness spending is particularly high for Nordic countries; for Denmark, Norway, Sweden, the Netherlands and Iceland, it exceeds 3% of GDP. These statistics do not include workers' compensation, private disability insurance or private sector spending. For Canada, it also does not include provincial level spending on social assistance. Consequently, comparability is an issue, particularly for countries with multiple programmes provided at different levels of government or distributed differently between the public and private sectors.

Most public disability spending is passive, i.e. in the form of benefits rather than employment (re)integration programmes. The latter is known as active programmes. Spending on such programmes is generally less than 8% of total public spending and in most cases less than 4% (OECD 2010). The exceptions, in terms of a higher proportion of spending on active programmes, are Norway, Denmark, the Netherlands, Sweden, Germany, Belgium and Poland, though for the last three countries it is low in terms of percentage of GDP.

Customarily, transfer payments are not included in burden calculations from the societal perspective because they are not a measure of resource consumption. Rather, they are simply a transfer of purchasing power from one group of individuals to another. They may be included in studies taking a disability system level perspective. Other times they might be included as a proxy measure for lost output. For the latter, they are poor approximations, since benefit levels in most disability programmes are substantially lower than the output loss associated with the disability.

Occupational injury and illness burden estimates produced by Leigh et al. (2001) provide a good example of how to estimate the burden at

Table 4.4 Total cost of occupational injuries and illnesses in California for 1992 (adapted from Leigh et al. 2001)

	Billions of dollars	Cost per incident
<i>1.65 M nonfatal injuries per year</i>		
<i>Direct costs</i>		
Medical costs	\$3.67	
Medical administration	\$0.81	
Indemnity administration	\$0.90	
<i>Total direct costs</i>	\$5.37	\$3,266
<i>Indirect costs</i>		
Cost of workplace training, re-staffing, disruption	\$0.29	
Lost earnings	\$8.66	
Lost fringe benefits	\$1.82	
Lost home production	\$1.15	
<i>Total indirect costs</i>	\$11.93	\$7,250
<i>1.33 M nonfatal illnesses per year</i>		
<i>Direct costs</i>		
Medical costs	\$0.47	
Medical administration	\$0.07	
Indemnity administration	\$0.03	
<i>Total direct costs</i>	\$0.56	\$422
<i>Indirect costs</i>		
Lost earnings	\$0.32	
Lost fringe benefits	\$0.06	
Lost home production	\$0.04	
<i>Total indirect costs</i>	\$0.42	\$313
<i>Overall total for nonfatal injuries and illnesses</i>	\$18.28	
<i>Overall total for fatal injuries and illnesses</i>	\$2.39	

the societal level. The Leigh et al. estimates are for California for the year 1992. The study considers both direct and indirect costs. Direct costs refer to medical expenses and insurance administration expenses (the latter does not include benefit expenses). Indirect costs refer to output losses consisting of lost earnings, fringe benefits and home production. The human capital approach is used to estimate output losses. The incidence-based approach is used, where the burden is based on lifetime costs of new cases arising in the calendar year. Table 4.4 provides summary measures for the direct and indirect costs.

The total costs burden for California was \$20.67 billion in 1992, with work disability costs (lost earnings and fringe benefits) from both non-fatal injuries and illnesses amounting to \$10.86 billion (approximately 50% of the total). These burden costs likely underestimate the true burden

because they do not consider the value of pain, suffering and loss of enjoyment of life, or home care provided by family members. The authors compare this burden to other health conditions such as AIDS, Alzheimer's disease and MSD conditions, and find that they are higher than each of them. The costs are similar to the cost of cancer, but slightly less than the cost of heart disease and stroke combined.

Using the same approach as above, Leigh et al. (1997) estimate the total burden of occupational injury and illness for the USA for calendar year 1992. In that year there were 13.2 million nonfatal injuries, 862,200 nonfatal diseases, 6,500 injury fatalities and 60,300 disease fatalities. The total direct cost across all categories for the year was \$65 billion and the indirect cost \$106 billion. An update on the US occupational injury and illness burden estimate for 2007

identified a total burden of \$246 billion, compared to an inflation-adjusted 1992 burden of \$217 billion (Leigh 2011). Other related work by Leigh et al. identifies the sectors in the USA with the top injury and illness costs in terms of average cost per worker (Leigh et al. 2004), and the states with the highest average (per worker) costs (Waehrer et al. 2004). The highest cost industries were taxicabs, bituminous coal and lignite mining, logging, crushing stone, oil field services, water transportation services, sand and gravel, and trucking. Southern and western states were more likely to be in the high cost per worker category, largely because of industry composition.

Based on Leigh et al. (2001), burden estimates for Canada were calculated for calendar year 2001 (Tomba 2002). In Canada, the direct cost of occupational injuries and illnesses exceeded \$6 billion per year. This estimate includes insurance administration expenses and medical services that are paid by employers through workers' compensation premiums. The indirect cost estimate for Canada is \$12 billion. This includes costs incurred by employers to accommodate injured workers who return to work, recruitment and training costs incurred for replacing injured workers, earnings lost by workers due to injury and the lost home production of workers. As with Leigh et al. (2001) these costs are likely an underestimate of the true societal burden, since they do not include costs associated with pain, suffering and loss of enjoyment of life or home care provided by family members. Furthermore, the number of claims is an underestimate of the true number of work-related injuries. Underreporting is well documented in the literature and is an issue that needs to be addressed if accurate estimates of burdens are to be calculated, since the magnitude of underreporting can be substantial (Shannon and Lowe 2002).

Estimating the burden of occupational injury and illness requires investigating multiple categories of costs associated with different stakeholders. What categories and what stakeholders will vary between countries due to differences in their disability policy systems. For example, in countries with comprehensive systems, such as

Germany, Austria and Switzerland, the direct costs are paid for by companies (Rauner et al. 2005), whereas in countries with low levels of social security such as the USA, individuals often bear a large fraction of the direct costs (Leigh et al. 2000). Rauner et al. (2005) divide the entities that might bear the burden of work disability into four broad categories: (1) social security, (2) private insurance companies, (3) employers and (4) others stakeholder such as individuals and society.

To estimate the output loss associated with work disability from all health conditions (whether work related or nonwork related) requires assumptions about the number and proportion of the disabled who would be working if not for their disability. This might be approximated by assuming levels of employment similar to their non-disabled counterparts. Wage rates attributable to labour-time loss also require approximation. Statistics on employment rates and average wage rates can be used to estimate these numbers.

EBIC (Health Canada 1998) estimates the output loss from all health conditions for Canada for calendar year 1998. The study uses a prevalence approach and considers both direct and indirect costs of morbidity and mortality. Direct costs in the study include medical care and rehabilitation costs. Indirect costs include lost earnings and home production. Table 4.5 provides details on the indirect costs estimated in this study. Overall the total burden for short- and long-term disability and premature mortality is \$75.5 billion, which amounts to 4.71% of GDP. The fraction attributable to work disability is 2.62% of GDP. This is a large burden and likely underestimates the true cost, since it only accounts for a few categories of costs.

4.9 Summary and Suggestions for the Way Forward

This chapter has provided an overview of the importance of burden studies and their value to the policy decision-making process. The chapter has also provided insights into the magnitude of

Table 4.5 Indirect costs of health conditions in Canada for 1998 (adapted from Health Canada 1998)

	Billions of dollars	Percentage of GDP	
<i>Indirect costs associated with short-term disability</i>			
Lost earnings	\$3.90	0.24	
Lost home production	\$5.90	0.37	
<i>Total indirect costs</i>	\$9.80		0.61
<i>Indirect costs associated with long-term disability</i>			
Lost earnings	\$13.00	0.81	
Lost home production	\$19.20	1.20	
<i>Total indirect costs</i>	\$32.20		2.01
<i>Indirect costs associated with premature mortality</i>			
Lost earnings	\$13.50	0.84	
Lost home production	\$20.00	1.25	
<i>Total indirect costs</i>	\$33.50		2.09
<i>Overall total for short-term and long-term disabilities and premature mortality</i>	\$75.50		4.71

the burden of work disability in a range of industrialized countries. The burden can be quite far reaching, encompassing working age adults whose engagement in the labour force is temporarily or permanently compromised due to a health condition. The burden includes both direct costs (i.e. health care, wage replacement benefits and rehabilitation services from various public/private insurance providers) and indirect costs (i.e. labour productivity/output losses). The full extent of the burden can and does encompass nonfinancial and sometime intangible outcomes such as individual role functioning outside of the paid labour force and health-related quality of life. Most studies on this topic have only been able to measure a part of the full burden. Even though they are not comprehensive in their accounting of burdens and costs, their findings suggest that the magnitudes are substantial. Given this fact, the issue of work disability and its prevention clearly warrants attention by policymakers, employers, labour representatives and society at large. There is much to be gained at the individual and societal level, particularly in terms of productivity and output for the economy, and also for functioning in other social roles. Once several interventions have been identified that appear appropriate for the context in which they are to be implemented, the obvious next step is to evaluate their effectiveness and cost-effectiveness before

they are adopted across the board. For that, we refer readers to Chap. 23 of this handbook, which provides guidance on the economic evaluation work disability prevention programmes.

References

- Bloom, D. E., & Canning, D. (2000). The health and wealth of nations. *Science's Compass*, 287, 1207–1209.
- Bloom, D. E., & Sachs, J. D. (1998). Geography, demography, and economic growth in Africa. *Brookings Papers on Economic Activity*, 29(2), 201–296.
- Brouwer, W. B. F., van Exel, N. J. A., Koopmanschap, M. A., & Rutten, F. F. (2002). Productivity costs before and after absence from work: As important as common? *Health Policy*, 61, 173–187.
- Health Canada. (1989). *Economic burden of illness in Canada*. Ottawa: Health Canada.
- Health Canada. (1996). *Economic burden of illness in Canada*. Ottawa: Health Canada.
- Health Canada. (1998). *Economic burden of illness in Canada*. Ottawa: Health Canada.
- de Jong, P. R., & de Vos, E. L. (2005). Lessons from the Dutch experience. *Revue Française des Affaires Sociales*, 2, 183–205.
- de Vos, K., Kapteyn, A., & Kalwij, A. (2012). Disability insurance and labor market exit routes of older workers in the Netherlands. In D. Wise (Ed.), *Social security programs and retirement around the world: Historical trends in mortality and health, employment, and disability insurance participation and reforms* (pp. 419–447). Chicago: Chicago University Press.
- Dembe, A. E. (2000). Pain, function, impairment, and disability: Implications for workers' compensation and

- other disability insurance systems. In T. G. Mayer, R. J. Gatchel, & P. B. Polatin (Eds.), *Occupational musculoskeletal disorders: Function, outcomes, and evidence* (pp. 563–576). Philadelphia: Lippincott, Williams & Wilkins.
- Department of Health and Human Services (DHHS). (2011). *Projected future growth of the older population*. Retrieved November 27, 2011, from http://www.aoa.gov/AoARoot/Aging_Statistics/future_growth/future_growth.aspx#age
- Drummond, M., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). Oxford: Oxford University Press.
- Grossman, M. (1972). *The demand for health: A theoretical and empirical investigation. Occasional Paper #119*. Cambridge, MA: National Bureau of Economic Research.
- Human Resources and Skills Development Canada (HRSDC). (2011). *Canadians in context - aging population*. Retrieved November 27, 2011, from <http://www4.hrsdc.gc.ca/3ndic.1t.4r@-eng.jsp?iid=33>
- Kessler, R. C., Greenberg, P. E., Mickelson, K. D., Meneades, L. M., & Wang, P. S. (2001). The effects of chronic medical conditions on work loss and work cut-back. *Journal of Occupational and Environmental Medicine, 43*(3), 218–225.
- Koopmanschap, M. A., Rutten, F. F. H., van Ineveld, B. M., & van Roijen, L. (1995). The friction cost method for measuring indirect costs of disease. *Journal of Health Economics, 14*(2), 171–189.
- Leigh, J. P. (2011). Economic burden of occupational injury and illness in the United States. *The Milbank Quarterly, 89*(4), 728–772.
- Leigh, J. P., Cone, J. E., & Harrison, R. (2001). Costs of occupational injuries and illnesses in California. *Preventive Medicine, 32*, 393–406.
- Leigh, J. P., Markowitz, S., Fahs, M., & Landrigan, P. (2000). *Costs of occupational injuries and illnesses*. Ann Arbor: The University of Michigan Press.
- Leigh, J. P., Markowitz, S. B., Fahs, M., Shin, C., & Landrigan, P. J. (1997). Occupational injury and illness in the United States: Estimates of costs, morbidity and mortality. *Archives of Internal Medicine, 157*(July 28), 1557–1568.
- Leigh, J. P., Waehrer, G., Miller, T. R., & Keenan, C. (2004). Costs of occupational injury and wellness across industries. *Scandinavian Journal of Work and Environmental Health, 30*(3), 199–205.
- Milligan, K., & Wise, D. A. (2012). Introduction and summary. In D. Wise (Ed.), *Social security and retirement around the world: Historical trends in mortality and health, employment, and disability insurance participation and reforms* (pp. 1–39). Chicago: Chicago University Press.
- Murray, C. J. L. (1994). Quantifying the burden of disease: The technical basis for disability-adjusted life years. *Bulletin of the World Health Organization, 72*(3), 429–445.
- Nagi, S. (1965). Some conceptual issues in disability and rehabilitation. In M. B. Sussman (Ed.), *Sociology and rehabilitation* (pp. 100–113). Washington: American Sociological Association.
- Nagi, S. (1991). Disability concepts revisited: Implications for prevention. In A. Pope & A. Tarlov (Eds.), *Disability in America: Towards a national agenda for prevention* (pp. 309–327). Washington: National Academy Press.
- Nicholson, S., Pauly, M. V., Polsky, D., Sharda, C., Szrek, H., & Berger, M. L. (2006). Measuring the effects of work loss on productivity with team production. *Health Economics, 15*, 111–123.
- OECD. (2010). *Sickness, disability and work: Breaking the barriers. A synthesis of findings across OECD countries*. Paris: OECD.
- Pauly, M. V., Nicholson, S., Xu, J., Polsky, D., Danzon, P. M., Murray, J. F., & Berger, M. L. (2002). A general model of the impact of absenteeism on employees. *Health Economics, 11*, 221–231.
- Peterson, M. A., Reville, R. T., & Stern, R. K. (1998). *Compensating permanent workplace injuries: A study of the California system*. Santa Monica: Rand Institute for Civil Justice.
- Polinder, S., Meerding, W. J., Mulder, S., Petridou, E., van Beeck, E., & EUROCCOST Reference Group. (2007). Assessing the burden of injury in six European countries. *Bulletin of the World Health Organization, 85*(1), 27–34.
- Rauner, M. S., Harper, P. R., & Schwarz, B. (2005). Economic impact of occupational injuries: Resource allocation for prevention programs. *Central European Journal of Operations Research, 13*(4), 393–424.
- Shannon, H. S., & Lowe, G. S. (2002). How many injured workers do not file claims for workers' compensation benefits? *American Journal of Industrial Medicine, 42*, 467–473.
- Sharpe, A., & Murray, A. (2010). *State of the evidence on health as a determinant of productivity*. Ottawa: Centre for the Study of Living Standards.
- Stewart, W. F., Ricci, J. A., Chee, E., & Morganstein, D. (2003). Lost productive work time costs from health conditions in the United States: Results from the American productivity audit. *Journal of Occupational and Environmental Medicine, 45*(12), 1234–1246.
- Tompa, E. (2002). The impact of health on productivity: Macroeconomic and microeconomic evidence and policy implications. In A. Sharpe, F. St-Hilaire, & K. Banting (Eds.), *The review of economic performance and social progress* (pp. 181–202). Montreal: Institute for Research on Public Policy.
- Tompa, E., de Oliveira, C., Dolinschi, R., & Irvin, E. (2008). A systematic review of disability management interventions with economic evaluations. *Journal of Occupational Rehabilitation, 18*(1), 16–26.
- Tompa, E., Scott-Marshall, H., Dolinschi, R., Trevithick, S., & Bhattacharyya, S. (2007). Precarious employment experiences and their health consequences: Towards a

- theoretical framework. *Work: A Journal of Prevention, Assessment and Rehabilitation*, 28(7), 209–224.
- Tompa, E., Scott-Marshall, H., Trevithick, S., & Bhattacharyya, S. (2006). Precarious employment and people with disabilities. In L. F. Vosko (Ed.), *Precarious employment in the Canadian labour market* (pp. 90–114). Montreal: McGill University Press.
- Waehrer, G., Leigh, J. P., Cassady, D., & Miller, T. R. (2004). Costs of occupational injury and illness across states. *Journal of Occupational and Environmental Medicine*, 46(10), 1084–1095.
- Weil, D. (2001). Valuing the economic consequences of work injury and illness: A comparison of methods and findings. *American Journal of Industrial Medicine*, 40, 418–437.
- World Health Organization. (1980). *International classification of impairments, disabilities and handicaps: A manual of classification relating to the consequences of disease*. Geneva: World Health Organization.
- World Health Organization. (2001). *International classification of functioning, disability and health*. Final draft, full version. Geneva: World Health Organization.
- World Health Organization. (2011). *Disability weights, discounting and age weighting of DALYs*. Retrieved November 12, 2011, from http://www.who.int/health-info/global_burden_disease/daly_disability_weight/en/index.html