Int J Public Health
DOI 10.1007/s00038-014-0553-y

HINTS AND KINKS

DALY calculation in practice: a stepwise approach

Brecht Devleesschauwer · Arie H. Havelaar · Charline Maertens de Noordhout · Juanita A. Haagsma · Nicolas Praet · Pierre Dorny · Luc Duchateau · Paul R. Torgerson · Herman Van Oyen · Niko Speybroeck

Received: 23 February 2013 / Revised: 22 March 2014 / Accepted: 3 April 2014

Introduction

The disability-adjusted life year (DALY) is an increasingly used population health metric (Murray et al. 2013; Devleesschauwer et al. 2014). DALYs are healthy life years lost, calculated by adding the adjusted number of years lived with disability (YLDs) and the number of years of life lost due to premature mortality (YLLs):

$$\text{YLD} = \text{Number of cases} \times \text{duration till remission or death} \times \text{disability weight} \quad (1)$$

$$\text{YLL} = \text{Number of deaths} \times \text{life expectancy at the age of death} \quad (2)$$

$$\text{DALY} = \text{YLD} + \text{YLL} \quad (3)$$

The philosophical and methodological aspects of the DALY calculation have been described (and debated) in great detail (Murray 1994; Murray and Acharya 1997; for critiques, see, amongst others, Anand and Hanson 1997), and are summarized in the accompanying Hints and Kinks paper (Devleesschauwer et al. 2014). The steps preceding the actual calculation, however, remain less well documented. This Hints and Kinks paper tries to address this gap by presenting a stepwise approach towards a DALY calculation.

Step 1: study population definition

As a first step, the context in which the burden assessment study will take place should be clearly defined. The target population must be delineated by defining study area and

B. Devleesschauwer (✉) · C. Maertens de Noordhout · N. Speybroeck
Faculty of Public Health, Institute of Health and Society (IRSS), Université Catholique de Louvain, Brussels, Belgium
e-mail: Brecht.Devleesschauwer@UGent.be

B. Devleesschauwer · P. Dorny
Department of Virology, Parasitology and Immunology, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium

A. H. Havelaar
Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

A. H. Havelaar
Institute for Risk Assessment Sciences, University Utrecht, Utrecht, The Netherlands

J. A. Haagsma
Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

N. Praet · P. Dorny
Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

L. Duchateau
Department of Comparative Physiology and Biometrics, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium

P. R. Torgerson
Section of Epidemiology, Vetsuisse Faculty, University of Zürich, Zurich, Switzerland

H. Van Oyen
Unit of Public Health and Surveillance, Scientific Institute of Public Health, Brussels, Belgium

Published online: 21 April 2014

time period. The latter may be one specific year, or a range of years, which can then be used to calculate the average burden of that time period.

Step 2: disease model definition

The disease model (also called outcome tree) serves as a guide through the further process of the study. Figure 1 presents the causal chain of disease. In general, risk factors increase the risk of disease, either directly or indirectly through facilitating exposure to biological, chemical or physical hazards. The course of disease is characterized by different health states (e.g., acute or chronic phases, short-term or long-term sequelae), possibly having different severity levels. A disease model is a schematic representation of the different health states associated with the concerned cause of disease burden, and the possible transitions between these states.

Depending on the cause of interest, i.e., the disease as such, the hazard or the risk factor, three different approaches may be distinguished:

1. *Outcome-based disease models* represent different health states of diseases, irrespective of the possible (infectious or non-infectious) aetiologies. For example, a disease model for the burden of diarrhoea could describe different diarrhoea severity levels (mild, moderate, severe), contributing YLDs, and diarrhoea-related death, contributing YLLs (Lamberti et al. 2012). Soerjomataram et al. (2012) present generic outcome-based disease models for cancer, including disease progression phases and sequelae.
2. *Hazard-based disease models* represent different health states associated with hazards such as biological or chemical agents or traumas (Mangen et al. 2013). For example, *Campylobacter* infection causes diarrhoea, but also other health states, such as Guillain–Barré syndrome (GBS) (Havelaar et al. 2000). The disease model would thus consist of the different associated symptoms (contributing YLDs), and death attributable to each symptom (contributing YLLs).
3. *Risk factor-based disease models* represent different health states associated with risk factors. For example, a disease model for unsafe water would include the health effects associated with feco-oral pathogens, including *Campylobacter*, other diarrhoeal pathogens and soil-transmitted helminths (Prüss et al. 2002).

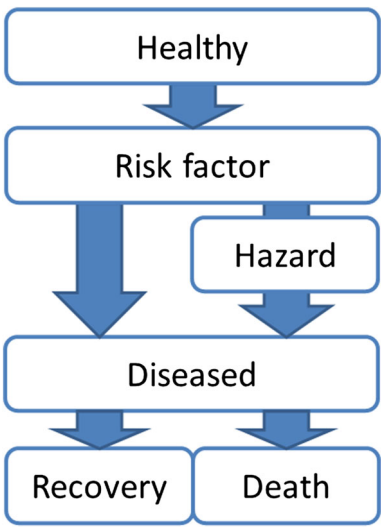

Fig. 1 Generic causal chain of disease

A hazard-based disease model for *Campylobacter* infection is given in Fig. 2. This model includes diarrhoea, possibly progressing from acute to chronic stages and to death; GBS occurring in a mild or severe form, the latter being able to cause death; and reactive arthritis. Other disease models can be constructed in a similar way. Disease models can be obtained from previous burden studies. However, new insights might indicate an updated disease model requiring data from a systematic review of relevant clinical and epidemiological aspects.

Step 3: data collection

This step is often the most difficult and time-consuming one. It is also the most crucial one, as the quality of the final DALY estimate directly depends on the quality of the data. Ideally, the necessary data should be collected through a systematic review of peer-reviewed literature and various sources of grey literature, including government agencies, non-governmental organizations and academia. An important guide for conducting and reporting systematic reviews is the PRISMA statement (Moher et al. 2009). As much as possible, collected data should be stratified by age and sex, as this will yield a more precise overall estimate and will enable to study the burden by age and sex. Further stratification by other parameters can also be useful and could allow a breakdown of disease burden by sub-region, occupation, socio-economic status, etc.

Extrapolation models may be needed when literature searches cannot provide essential data. These models estimate parameters from data of neighbouring regions or other time periods. The external data used must thus be representative of the selected population, region and time. Where no empirical data can be identified, expert elicitation may be applied (see Cooke 1991 for a guide).

In general, three groups of data need to be collected.

Demographical data

DALY calculations require the total number of males and females, per age group, of the selected area and time

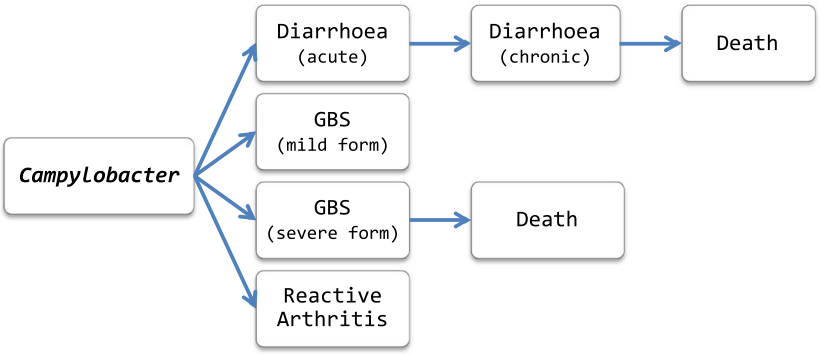

Fig. 2 Hazard-based disease model for *Campylobacter* infection (*GBS* Guillain–Barré syndrome)

period. These data can be obtained from national statistical bureaux or from the United Nations Statistics Division (available from http://data.un.org/). A second required source of demographical data is the life-expectancy table. In order to enhance comparability, the life expectancy at the age of death is generally drawn from standard life-expectancy tables (Devleesschauwer et al. 2014). However, local life-expectancy tables may also be used, in particular when within-country disease ranking is an important objective. Local life tables are typically available from national statistical bureaux.

Epidemiological data

The most important data for DALY calculations are the number of cases in the different health states defined by the disease model, including death. Most often, the number of incident cases is considered, although DALYs may also be calculated based on the number of prevalent cases (see Schroeder 2012 for a comparison). In the following, we will assume an incidence perspective.

The number of incident cases is the product of the incidence rate and the population size. Depending on data availability, there are three approaches to obtain incidences of the individual health states in the disease model:

1. *Direct approach* The incidence of a health state is directly available, e.g., through a disease or mortality register. In our example (Fig. 1), this would be the case if a prospective population-based study directly estimated the incidence of *Campylobacter*-associated GBS.
2. *Attribution approach* The incidence of a health state is obtained from the overall incidence of the given health state (i.e., regardless of aetiology), and an attribution probability (not applicable for outcome-based disease models):

Overall incidence health state
× proportion attributable to hazard or risk factor

The incidence of GBS due to *Campylobacter* can for instance be obtained by multiplying the overall incidence of GBS in the population to the proportion of *Campylobacter*-attributable GBS. In risk factor-based disease models, the proportion of cases attributable to the risk factor is commonly referred to as the population attributable fraction (Perez and Künzli 2009).

3. *Transition approach* The incidence of a health state can be obtained from the previous incidence in the model. Two scenarios are possible:

Overall incidence burden cause × probability of transition to health state
Incidence health state × probability of transition to next health state

The first scenario would imply that we model the incidence of *Campylobacter*-associated GBS on the overall *Campylobacter* incidence in the population and the probability of developing GBS after *Campylobacter* infection. The second scenario would imply, for instance, that we model the chronic diarrhoea incidence by multiplying the acute diarrhoea incidence by the probability of progressing from acute to chronic diarrhoea.

Possibly, epidemiological estimates can be derived through different approaches. In this case, cross-validation of the different estimates can be performed, which can strengthen their reliability.

Disease severity data

DALY estimates incorporate the severity of the health states through their duration and disability weight (DW)

[see Eq. (1)]. The former can be obtained through hospital registers, literature reviews or expert elicitation. The latter are most commonly derived from global burden of disease studies (Murray and Lopez 1996; Salomon et al. 2013) or the Dutch burden of disease study (Stouthard et al. 2000). Essink-Bot and Bonsel (2002) describe methods for DW derivation.

Step 4: data adjustment

Potential data biases should be critically appraised. Under-reporting and under-ascertainment are well-documented sources of bias (Lake et al. 2010). Misclassification bias due to imperfect diagnostic tests can be amended using various statistical techniques (Speybroeck et al. 2013). Finally, coherence in different epidemiological parameters for non-infectious diseases can be assessed using DisMod software (Barendregt et al. 2003).

Step 5: DALY calculation

Once all required data have been gathered, the actual DALY calculation can commence. DALYs for a disease, hazard or risk factor are obtained by summing the YLDs and YLLs calculated for each health state in the disease model. For a technical summary of DALY calculations, see Devleesschauwer et al. (2014). Despite wide availability of DALY calculation methods, standardized DALY calculation tools are currently not readily accessible. The World Health Organization provides a template for deterministic DALY calculations (available from http://www.who.int/entity/healthinfo/bodreferencedalycalculationtemplate.xls). Alternatively, the DALY calculator in R can be used for stochastic DALY calculations (Devleesschauwer et al. 2013). Until now, however, most users have made their own calculation model, either in a programming language such as R, or in Microsoft Excel with add-ins enabling Monte Carlo simulation. Although flexible, these methods do not guarantee methodological transparency. Further standardization of DALY calculation tools, including databases with default datasets, is therefore necessary to improve consistency in DALY calculations and to further DALY-based burden of disease studies.

References

Anand S, Hanson K (1997) Disability-adjusted life years: a critical review. J Health Econ 16:685–702

Barendregt JJ, Van Oortmarssen GJ, Vos T, Murray CJ (2003) A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. Popul Health Metr 1:4

Cooke R (1991) Experts in uncertainty: opinion and subjective probability in science. Oxford University Press, New York

Devleesschauwer B, McDonald S, Haagsma J, Praet N, Havelaar A, Speybroeck N (2013) DALY: DALY calculator—a GUI for stochastic DALY calculation in R. R package version 1.2.0. http://CRAN.R-project.org/package=DALY

Devleesschauwer B, Havelaar AH, Maertens de Noordhout C, Haagsma JA, Praet N, Dorny P, Duchateau L, Torgerson PR, Van Oyen H, Speybroeck N (2014) Calculating disability-adjusted life years to quantify burden of disease. Int J Public Health. doi:10.1007/s00038-014-0552-z

Essink-Bot ML, Bonsel GJ (2002) How to derive disability weights? In: Murray CJ, Lopez AD, Salomon JA (eds) Summary measures of population health: concepts, ethics, measurement and applications. World Health Organization, Geneva, pp 449–465

Havelaar AH, de Wit MA, van Koningsveld R, van Kempen E (2000) Health burden in the Netherlands due to infection with thermophilic *Campylobacter* spp. Epidemiol Infect 125:505–522

Lake RJ, Adlam SB, Perera S, Campbell DM, Baker MG (2010) The disease pyramid for acute gastrointestinal illness in New Zealand. Epidemiol Infect 138:1468–1471

Lamberti LM, Fischer Walker CL, Black RE (2012) Systematic review of diarrhea duration and severity in children and adults in low- and middle-income countries. BMC Public Health 12:276

Mangen MJ, Plass D, Havelaar AH, Gibbons CL, Cassini A, Mühlberger N, van Lier A, Haagsma JA, Brooke RJ, Lai T, de Waure C, Kramarz P, Kretzschmar ME, BCoDE consortium (2013) The pathogen- and incidence-based DALY approach: an appropriate [corrected] methodology for estimating the burden of infectious diseases. PLoS One 8:e79740

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6:e1000097

Murray CJL (1994) Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ 72:429–445

Murray CJL, Acharya AK (1997) Understanding DALYs. J Health Econ 16:703–730

Murray CJL, Lopez AD (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. Harvard University Press, Cambridge

Murray CJ, Vos T, Lozano R et al (2013) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the global burden of disease study 2010. Lancet 380:2197–2223

Perez L, Künzli N (2009) From measures of effects to measures of potential impact. Int J Public Health 54:45–48

Prüss A, Kay D, Fewtrell L, Bartram J (2002) Estimating the burden of disease from water, sanitation, and hygiene at a global level. Environ Health Perspect 110:537–542

Salomon JA, Vos T, Hogan DR et al (2013) Common values in assessing health outcomes from disease and injury: disability weights measurement study for the global burden of disease study 2010. Lancet 380:2129–2143

Schroeder SA (2012) Incidence, prevalence, and hybrid approaches to calculating disability-adjusted life years. Popul Health Metr 10:19

Soerjomataram I, Lortet-Tieulent J, Ferlay J, Forman D, Mathers C, Parkin DM, Bray F (2012) Estimating and validating disability-adjusted life years at the global level: a methodological framework for cancer. BMC Med Res Methodol 12:125

Speybroeck N, Devleesschauwer B, Joseph L, Berkvens D (2013) Misclassification errors in prevalence estimation: bayesian handling with care. Int J Public Health 58:791–795

Stouthard MEA, Essink-Bot M-L, Bonsel GJ, on behalf of the Dutch Disability Weights Group (2000) Disability weights for diseases: a modified protocol and results for a Western European region. Eur J Public Health 10:24–30