Disability Adjusted Life Years and Acute Onset Disorders

Improving estimates of the non-fatal burden of injuries and infectious intestinal disease

Disability adjusted life years en acute aandoeningen

Verbetering van ziektelastschattingen van ongevallen en maagdarminfecties

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus
Prof.dr. H.G. Schmidt
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 3 november 2010 om 13.30 uur

door

Juanita Antje Haagsma

geboren te Heerenveen

Promotiecommissie

Promotor:	Prof.dr. G.J. Bonsel
Overige leden:	Prof.dr. J.J. van Busschbach Prof.dr. A.H. Havelaar Prof.dr. P. Patka
Co-promotor:	Dr. E.F. van Beeck

Colofon

ISBN 978-90-8559-094-1

Lay-out: Juanita A. Haagsma
Cover illustration: Michael Johnson
Printed by: Optima Grafische Communicatie b.v.
This thesis was printed with the financial support of the institute of Health Policy and Management, Erasmus University Rotterdam.

Contents

1

Introduction

Summary measures of population health

The population's health faces an array of diseases and injuries. Limited resources compel policy-makers everywhere to focus on threats that are regarded most relevant in terms of public health. The verdict 'relevant in terms of public health' preferably rests on an objective and transparent tool. Researchers of the World Health Organization (WHO) and Worldbank developed an innovative concept called the burden of disease, which quantifies the health status of a population in comprehensive measures to aid policy-makers in science and resource allocation. Burden of disease studies provide knowledge on the size of health problems and the potential benefit of proposed measures set against similar and comparable data of other problems [1, 2].

Previously, cause specific mortality statistics and, to a certain degree, incidence data on predominantly infectious diseases and injuries have been used routinely to indicate the burden of these diseases. Impacts were then expressed as incidence changes. The shortcomings are obvious: long-term sequelae can not be deducted from incidence and diseases show great variety in severity, duration, and consequent treatment costs.

With increasing life span, information on long-term morbidity and disability has become more important. Furthermore, the use of methods of economic evaluation calls for uniform measures that integrate information regarding mortality, morbidity and disability on the individual level, with a lifetime-perspective rather than focussed on the date of contracting the disease [2, 3].

The combined demands of public health and health economics induced the development of summary measures of population health, that combined mortality and life expectancy on the one hand and morbidity and disability on the other hand. This thesis is dedicated to these measures, which first have been put forward in the public health domain by researchers of WHO and Worldbank creating the first burden of disease study published in 1993 [4]. In particular, this thesis is dedicated to the improvement of what appears to be the most significant step forward in public health methodology since the development of multi-factorial risk models for chronic disease.

Health-adjusted life years

Summary measures of population health can be generally labeled as health-adjusted life years (HALYs). The key feature of HALYs is that they integrate the impact of mortality along with morbidity by translating morbidity into a theoretical survival loss of assumed equivalent impact [2, 5, 6]. This allows summation of mortality and morbidity impact, and subsequent comparison of disease impact on public health across a range of illnesses and populations. The most popular HALYs are Quality Adjusted Life Years (QALYs) and Disability Adjusted Life Years (DALYs).

Quality Adjusted Life Years

QALYs provide a vital tool for economic evaluations, which have a strict format derived from welfare economics; these evaluations compare the costs and benefits of alternative health care interventions [7-9]. The principal economic evaluation technique is cost-utility analysis, which compares the costs, measured in monetary units, with the benefits, expressed in improvements in survival and health-related quality of life, the latter expressed in a value score between 0 and 1 [10]. Essentially, QALYs represent survival that is down-weighted for the time lived with functional capacity.
In hindsight, the paper of Williams represents a landmark paper in the application of QALYs to discuss policy options in cardiovascular disease treatment [11]. From then on, QALYs conquered the world of intervention evaluation research. Despite the fact that the concept rested on many assumptions and arbitrary choices, its face validity paved the way. The current state of play is that QALYs are the principal choice to express and compare pharmaceutical benefits; they are indispensable in search of the value for money of expensive new drugs.

Disability Adjusted Life Years

On a parallel track, the DALY concept was developed and subsequently used in the ground-breaking 1993 Development Report and 1996 Global Burden of Disease (GBD) study [4, 12]. Within years the DALY radically changed population epidemiology and health policy at the macro level. It has since grown into the standard outcome measure in public health and the primary outcome for priority-setting in health care and prevention.
The DALY measures the burden of disease, i.e. it aggregates the total health loss at population level into a single index by summarizing a) years of life lost (YLL), and b) years lived with disability (YLD) [13]. YLL represents the time lost through premature mortality and YLD represents the healthy time lost while living with a disability. The DALY aggregates mortality, expressed in years of life lost (YLL) and morbidity, expressed in years lived with disability (YLD) into a single figure, and is calculated as:

DALY = YLL + YLD.

YLL represents the time lost due to premature mortality and is calculated with the following formula:

$$YLL = \Sigma\, d_l * e_{l,}$$

where d are the number of fatal cases due to health outcome l in a certain period and e is the expected individual life span at the age of death.

YLD represents the healthy time lost while living with a disease or disability and is calculated with the following formula:

$$YLD = \Sigma n_l * t_l * dw_l,$$

where n are the number of cases with health outcome l, t the duration of the health outcome and dw the disability weight assigned to health outcome l.

QALY and DALY

QALYs and DALYs have much in common. Both measures integrate mortality and morbidity. To measure morbidity, survival is weighted for the time lived with functional capacity. This functional capacity is expressed in health-related quality of life and captured on a scale with endpoints 0 and 1. These endpoints represent the extremes worst possible health state and full health. On this scale, disease are assigned a value, or weight, that reflects the effect of the disease on health-related quality of life. To produce estimates of DALYs or QALYs, life expectancy is then multiplied by weights associated with different diseases.

Disability weights

As mentioned above, the disability weight is an essential factor for establishing DALYs (and QALYs). The disability weight is a value that is assigned to living with disability and this value is commonly based on elicited preferences obtained from a panel of judges [14, 15]. Preferences are defined as quantitative expressions or valuations for certain health states, which reflect the relative desirability of the health states [7, 16]. Disability weights are subject to design choices which affect the resulting values. The main choices concern: (1) the panel of judges, (2) the valuation method and (3) the generic or disease-specific method to depict the disease. Whether one expresses the weight as 'gain' (QALY) or 'loss' (DALY) is trivial from the conceptual point of view.

Disability weights and the panel

The panel of judges providing the preferences may consist of patients or a valid proxy, medical experts or members of the general public [17-19] . Overall, preferences derived from patients are lower compared to preferences derived from the general public, resulting in higher disability weights [20-22], though these findings are ambiguous [23, 24]. It has been recommended to use disability weights which are based on societal preferences, because burden of disease studies are primarily used as a tool for guiding decision-making on resource allocation at the population level [2].

Disability weights and the valuation method

To measure individual preferences, several measurement methods exist, of which the visual analogue scale (VAS) and the time trade-off (TTO) have been applied most extensively [7, 16, 17]. Other methods like the person trade-off and standard gamble are

much less used, for two reasons: the are cumbersome and their initial attractiveness from a theoretical point of view did not hold. Each measurement tool affects the preference outcomes in a method-specific way, limiting exchange and comparability of preference data from different sources [25, 26].

Disability weights and generic versus disease specific presentation

In addition to the panel composition and the valuation method used, a third methodological aspect related to the disability weight is the approach chosen to depict the disease [14, 16, 27]. In the generic approach, the disability weight is assessed with a multi-attribute utility instrument (MAUI) and in the second approach, disease-specific descriptions are valued to derive disability weights

Generic presentation

With MAUI, generic attributes are used to classify health states [20, 28, 29]. Firstly patients describe their health state by choosing a functional level for each attribute. Using weights for the separate attributes, the reported functional level on the attributes is then converted into a disability weight which by definition fits within the 0-1 range. The weights that are used to convert the health states into a disability weight are derived at an earlier stage and they are based on preference data of the general population for health states described with the generic attributes. Widely used MAUIs are for instance the EQ-5D and Health Utilities Index (HUI) [30, 31]. Regarding the EQ-5D, the health profile is rendered by five attributes (mobility, self care, usual activities, pain/discomfort and anxiety/depression) and the tariff to calculate utilities consists of attribute weights that were derived with the TTO technique [32]. The HUI Mark 3 consists of attributes vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain and the attribute weights were derived with the Standard Gamble technique [33].
Important to note is that all MAUIs are generic instruments: both the functional health state of the patient and the tariff weights derived from the population are based on generic attributes and without regard to the underlying condition. Generic health state descriptions provide disability weights for a great number of possible generic health profiles; 243 in case of EQ-5D through a staggering 972,000 in case of HUI3. A generic approach therefore relies on the capacity to cover all relevant features of a disease.

Disease specific presentation

Another approach to assess disability weights is the valuation of disease-specific descriptions by a panel. With this approach disability weights are derived from disease specific health state valuations of a panel of judges. Disease specific health state descriptions indicate the cause and the specific health effects of the condition. They provide information that is not reflected in the generic health states, but which matters for health state valuation [34]. For instance, conditions with a bad prognosis yield

significantly higher disability weights when presented together with additional disease specific information rather than presented through generic information only. The panel study approach can be distinguished further into period profiles and annual health profiles. Period profiles assume that the health state remains constant over time. The panel study approach with period profiles has been applied in the GBD 1996 study [12, 13, 15]. With the annual profile approach, the assumption of separated states of constant health is released and the health profile is valued as a whole, alleviating the contrived assumption of constancy of health. This allows valuation of health profiles characterized by dynamic and/or complex patterns, such as acute non-stable health profiles [14]. The annual profile approach was first applied in the Dutch Disability Weights study [34] and further investigated in the Mild diseases and ailments (MiDAS) study [35].

Methodological constraints of the global burden of disease study

The groundbreaking 1996 GBD study aimed to estimate the total burden of disease worldwide as the sum of all separate diseases [12]. The results of this endeavour emphasized more than ever the inequalities that exist between countries and regions around the globe. However, by virtue of its success, this ambitious project raised concerns regarding some of its methodological choices. This thesis addresses four of these controversial aspects, namely (1) deriving disability weights for disorders with complex and heterogeneous recovery patterns, (2) the disregard of comorbid diseases, (3) the arbitrariness of the in- and exclusion of long-term sequelae, and (4) the absence of a criterion to identify cases that are relevant from a public health perspective.

DALY in case of complex and heterogeneous patterns over time

For the GBD study a large set of disability weights had to be derived [12, 36]. Over one hundred specific disease categories with a disability component have been distinguished.
The initial approach used by Murray et al. was based on choices that have subsequently been challenged [15, 36]. Firstly, to arrive at the GBD disability weights, the disability-causing health states were presented to the panel of judges as a limited diagnostic description of the disease, derived from the terminology of the International Classification of Diseases (ICD). Apart from the disadvantage that only physicians could be involved in the evaluation, later studies showed that, even for expert panels, adding a generic description of functional health status to the diagnostic disease label is necessary to standardize the stimulus [34]. Secondly, in computing the DALYs, the GBD approach assumes independence between duration and disability and requires that the health state remains constant over time. As mentioned above, this is an assumption that is untenable for disorders that are characterized by complex time-severity [35].
A solution to this problem might be to derive and apply disability weights with the annual profile approach for such disorders. By describing the health profile over one year time, the annual profile approach overcomes the constraints of the MAUI and period profile

disability weights, allowing the derivation of more valid disability weights for acute onset disorders, episodic diseases, and health states characterized by complex and heterogeneous recovery patterns [14].

DALY in case of comorbidity

A second weakness concerns the disregard of comorbidity [15, 37]. Comorbidity here is defined as the presence of any clinical condition which qualifies for formal classification as a disease, different from the disease under study. It can be divided into at least three types: unrelated (the most common case of two conditions happening by chance on the same individual), related through common risk factors (pathophysiology is grossly unrelated, in particular the pathway to symptoms), and directly related where one condition can be regarded as natural consequences or parallel manifestation of the other condition.

Up till now the majority of burden of disease studies have failed to consider the fact that more than one condition may exist simultaneously in patients. This neglect has implications for priority setting and prevention, particularly if the primary disease complicates the prognosis of other diseases. Also, one easily overestimates the gains from care (e.g. prevention or therapy) if a substantial number of patients suffers from additional conditions. As a result, ignoring the impact of comorbid conditions yields wrongful policy recommendations. Counting all co-existing conditions as separate entities leads to a gross overestimation of the ill health status in the population due to double counting. Moreover, in an epidemiological model, it is frequently rather difficult to cover the multiple pathways of mutual influence of co-existing clinical conditions. Perhaps as a consequence of this complexity, so far little attention has been directed towards estimating disability weights for comorbid conditions.

The few studies that have adjusted for the impact of comorbidity applied an adjustment approach that involved the disability weight [38-40]. An example of such an adjustment approach is the multiplicative adjustment approach, which assumes that the additional effect of a comorbid disease increases the disability weight, though it is less than the sum of the disability weights of the diseases separately. Empirical findings based on actual patient data to verify the validity of the proposed comorbidity adjustment approaches are virtually nonexistent.

DALY in case of long-term sequelae

Another issue of accurate and valid assessment of burden of disease concerns the coverage of long-term health outcomes or sequelae. A sequela is a pathological condition resulting from a prior disease, injury, or attack. If burden of disease estimates aim to support policy makers, the assessments of health losses have to be complete, even if some of these are remote. All sequelae of a defined cause of ill health or disease should be included in the burden of disease estimate. Such being the case, the association between a prior disease and a sequela will not invariably be obvious. The time delay can cast doubt

on the strength of the causal relation, and in other cases the association between a disease and its sequelae has recently come to light or has been disputed. In some cases, the sequel is regarded independent disease, in particular if lag time is long, and pathophysiology and symptoms are quite different from the initial disease (e.g. hepatitis B and the sequel liver cell carcinoma). The ignorance of sequelae is more striking in other cases where well-known, undisputed, sequelae are ignored in DALY computations, apparently without obvious reason. For instance, burden of injury estimates usually focus on the physical consequences of an injury (i.e. hip fracture or intracranial injury), whereas psychopathological consequences such as posttraumatic stress disorder (PTSD), travel anxiety and other maladaptive syndromes are not considered. In the case of injury sequelae, time delay between exposure and consequence can only be a partial reason for this neglect, as many studies have revealed high prevalence of these disabling consequences among injury victims relatively shortly after the injury [41-45]. Perhaps the distinct nature of consequences plays a role. Nevertheless, the high impact on health-related quality of life and the usually long term presence of psychiatric consequences amount to a substantial number of DALYs [46, 47]. To satisfy the aspiration of burden of disease studies, namely identifying priorities in preventive efforts and medical resource allocation, such sequelae should be taken into account even if their measurement is demanding.

DALY in case of minimal disease

Related to the question regarding the boundaries of associated consequences in terms of the permissible time delay of their occurrence and certainty of their causal relation, is the question whether there should be a threshold, a cut-off point, reflecting a relevance criterion, which decides on the minimum of health impact accounted for in the calculation of any YLDs at all. Below an assumed threshold one might regard consequences as part of normal life, i.e. on average too trivial to account for. The DALY calculus allows adding many trivial health impacts to one major event: e.g. one hundred bruises may add to one fracture. In particular in the health domains that are featured by heterogeneous health outcomes and that include the extremes of the severity spectrum, such as the domain of injury, the concept of a maximal tolerable, zero-disability threshold on the individual level would influence the overall rankings according to population impact.

The application of burden of disease estimates in prioritization discussions will change - in our view improve - if a criterion to distinguish relevant disease from trivial disturbance is included in the burden of disease calculations. Several options for such a relevance criterion are available. It may be based on health outcome derivatives, health care use, or - in case of injuries - anatomical features[48]. A prerequisite, however, is that the criterion should be unambiguous, and still allow relevant mild disease to be included in burden of disease estimates.

Objective and outline of this thesis

The main objective of this thesis was to improve estimates the methodology of the DALY measure by addressing questions concerning key aspects of the methodology that have been raised. Alternative solutions for these controversial methodological aspects have been tested in two health (care) domains where these methodological aspects are particularly problematic, i.e. injuries and intestinal infectious disease. These health domains have in common that they are featured by heterogeneous health outcomes, including the extremes of the severity spectrum and duration and with all types of time-severity relations. As a result, these health domains in particular meet the methodological constraints of the DALY measure.

The objective of this thesis was operationalized in the following research questions:

- How can the assessment of health-related quality of life and the process to derive disability weights regarding acute onset disorders with a highly variable course over time be improved?
- Which of three existing comorbidity approaches in DALY calculations performs best?
- What is the impact of commonly ignored long-term sequelae of acute onset disorders on burden of disease estimates?
- How can a threshold of triviality be developed to separate cases of acute onset disorders which are relevant for public health policy from those which can validly be ignored regardless of their number (prevalence)?

Research question 1: How can the assessment of health-related quality of life and the process to derive disability weights regarding acute onset disorders with a highly variable course over time be improved?

Part I (Chapter 2 to 5) addresses the limitations of current methods regarding measurement of health-related quality of life and the assessment of the disability weight for disorders with complex and heterogeneous patterns over time and explores the performance of alternative methods.
Chapter 2 presents the development and application of novel APM disability weights for injury consequences. The APM disability weights are compared to MAUI disability derived patient survey data in Chapter 3. In Chapter 4 disease specific and generic disability weights derived with EQ-5D are compared. Health-related quality of life among injury patients measured with two MAU instruments are compared in Chapter 5.

Research question 2 Which of three existing comorbidity approaches in DALY calculations performs best?

Part II (Chapter 6) presents a systematic comparison of three comorbidity adjustment approaches in patients with an injury condition as primary condition and common conditions with non-trivial health impact as the secondary condition.

Research question 3: What is the impact of commonly ignored long-term sequelae of acute onset disorders on burden of disease estimates?

Part III (Chapter 7 to 9) addresses inaccuracies in YLD calculations that are caused by ignoring marked long-term sequelae in burden of disease estimates. The associating between the acute onset disorder and ignored long-term sequela is examined and the impact of including those sequelae in burden of disease estimates is assessed.
In Chapter 7 the association and effect of the inclusion of the sequela post-infectious irritable bowel syndrome in the burden of infectious intestinal disease is explored. Chapter 8 and 9 address the association between PTSD and injury. Chapter 8 presents the prevalence rate and the predictors of PTSD among injury victims and describes the impact on health-related quality of life. Chapter 9 explores the effect of including PTSD in the burden of injury calculations.

Research question 4: How can a threshold of triviality be developed to separate cases of minimal disease which are relevant for public health policy from those which can validly be ignored regardless their number?

Part IV (Chapter 10 to 11) focuses on the development of a threshold of triviality to identify, measure and account for disorders that are relevant for public health policy.
In Chapter 10 the performance of indicators based on health care, anatomical and outcome criteria are explored. Chapter 11 presents the application of a novel relevance criterion that is based on preferences to rank enteric pathogens and that may be used as a threshold of triviality.

References

1. Murray, C.J.L., A.D. Lopez, and C.D. Mathers, Summary measures of population health: concepts, ethics, measurement and applications. 2002, Geneva: World Health Organization.
2. Field, M.J. and M.R. Gold, Summarising population health: Directions for the development and application of population health metrics. 1998, Washington D.C.: Institute of Medicine: National Academy Press.
3. Etches, V., et al., Measuring population health: a review of indicators. Annu Rev Public Health, 2006. 27: p. 29-55.
4. Worldbank, World Development Report 1993: Investing in Health. 1993, New York: Oxford University Press.
5. Hyder, A.A., G. Rotllant, and R.H. Morrow, Measuring the burden of disease: healthy life-years. Am J Public Health, 1998. 88(2): p. 196-202.

6. Murray, C.J. and A.D. Lopez, Global comparative assessments in the health sector. 1994, Geneva: World Health Organization.
7. Drummond, M., et al., Methods for the Economic Evaluation of Health Care Programmes. 1997, Oxford, U.K.: Oxford Medical Publications.
8. Torrance, G.W., W.H. Thomas, and D.L. Sackett, A utility maximization model for evaluation of health care programs. Health Serv Res, 1972. 7(2): p. 118-33.
9. Gold, M.R., et al., Cost-effectiveness in health and medicine. 1996, New York: Oxford University Press.
10. Torrance, G.W., Measurement of health state utilities for economic appraisal. J Health Econ, 1986. 5(1): p. 1-30.
11. Williams, A., Economics of coronary artery bypass grafting. Br Med J (Clin Res Ed), 1985. 291(6491): p. 326-9.
12. Murray, C.J.L. and A.D. Lopez, The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. 1996, Cambridge: Harvard University Press.
13. Murray, C.J., Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ, 1994. 72(3): p. 429-45.
14. Essink-Bot, M.L. and G.J. Bonsel, How to derive disability weights?, in Summary measures of population health: concepts, ethics, measurement and applications., C.J.L. Murray, A.D. Lopez, and J.A. Salomon, Editors. 2002, World Health Organization: Geneva.
15. Anand, S. and K. Hanson, Disability-adjusted life years: a critical review. J Health Econ, 1997. 16(6): p. 685-702.
16. Froberg, D.G. and R.L. Kane, Methodology for measuring health-state preferences--I: Measurement strategies. J Clin Epidemiol, 1989. 42(4): p. 345-54.
17. Brazier, J., M. Deverill, and C. Green, A review of the use of health status measures in economic evaluation. J Health Serv Res Policy, 1999. 4(3): p. 174-84.
18. Dolan, P., et al., An inquiry into the different perspectives that can be used when eliciting preferences in health. Health Econ, 2003. 12(7): p. 545-51.
19. De Wit, G.A., J.J. Busschbach, and F.T. De Charro, Sensitivity and perspective in the valuation of health status: whose values count? Health Econ, 2000. 9(2): p. 109-26.
20. McNamee, P. and J. Seymour, Comparing generic preference-based health-related quality-of-life measures: advancing the research agenda. Expert Rev Pharmacoecon Outcomes Res, 2005. 5(5): p. 567-81.
21. Boyd, N.F., et al., Whose utilities for decision analysis? Med Decis Making, 1990. 10(1): p. 58-67.
22. Sackett, D.L. and G.W. Torrance, The utility of different health states as perceived by the general public. J Chronic Dis, 1978. 31(11): p. 697-704.
23. Balaban, D.J., et al., Weights for scoring the quality of well-being instrument among rheumatoid arthritics. A comparison to general population weights. Med Care, 1986. 24(11): p. 973-80.
24. Dolders, M.G., et al., A meta-analysis demonstrates no significant differences between patient and population preferences. J Clin Epidemiol, 2006. 59(7): p. 653-64.
25. Green, C., J. Brazier, and M. Deverill, Valuing health-related quality of life. A review of health state valuation techniques. Pharmacoeconomics, 2000. 17(2): p. 151-65.
26. Krabbe, P.F., M.L. Essink-Bot, and G.J. Bonsel, The comparability and reliability of five health-state valuation methods. Soc Sci Med, 1997. 45(11): p. 1641-52.
27. Tsuchiya, A. and P. Dolan, The QALY model and individual preferences for health states and health profiles over time: a systematic review of the literature. Med Decis Making, 2005. 25(4): p. 460-7.
28. Hawthorne, G. and J. Richardson, Measuring the value of program outcomes: a review of multiattribute utility measures. Expert Rev Pharmacoecon Outcomes Res, 2001. 1(2): p. 215-28.
29. Essink-Bot, M.L., et al., An empirical comparison of four generic health status measures. The Nottingham Health Profile, the Medical Outcomes Study 36-item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. Med Care, 1997. 35(5): p. 522-37.

30. Feeny, D., et al., Multi-attribute health status classification systems. Health Utilities Index. Pharmacoeconomics, 1995. 7(6): p. 490-502.
31. Rabin, R. and F. de Charro, EQ-5D: a measure of health status from the EuroQol Group. Ann Med, 2001. 33(5): p. 337-43.
32. Dolan, P., Modeling valuations for EuroQol health states. Med Care, 1997. 35(11): p. 1095-108.
33. Horsman, J., et al., The Health Utilities Index (HUI): concepts, measurement properties and applications. Health Qual Life Outcomes, 2003. 1: p. 54.
34. Stouthard, M.E., M.L. Essink-Bot, and G.J. Bonsel, Disability weights for diseases. A modified protocol and results for a Western European region. European Journal of Public Health, 2000. 10: p. 24-30.
35. Janssen, M.F., E. Birnie, and G. Bonsel, Feasibility and reliability of the annual profile method for deriving QALYs for short-term health conditions. Med Decis Making, 2008. 28(4): p. 500-10.
36. Murray, C.J. and A.K. Acharya, Understanding DALYs (disability-adjusted life years). J Health Econ, 1997. 16(6): p. 703-30.
37. Mathers, C., T. Vos, and C. Stevenson, The burden of disease and injury in Australia. 1999, Australian Institute of Health and Welfare: Canberra.
38. Mathers, C.D., K.M. Iburg, and S. Begg, Adjusting for dependent comorbidity in the calculation of healthy life expectancy. Popul Health Metr, 2006. 4: p. 4.
39. van Baal, P.H., et al., Disability weights for comorbidity and their influence on health-adjusted life expectancy. Popul Health Metr, 2006. 4: p. 1.
40. Mathers, C.D., et al., The burden of disease and injury in Australia. Bull World Health Organ, 2001. 79(11): p. 1076-84.
41. Ehlers, A., R.A. Mayou, and B. Bryant, Psychological predictors of chronic posttraumatic stress disorder after motor vehicle accidents. J Abnorm Psychol, 1998. 107(3): p. 508-19.
42. Ellis, A., G. Stores, and R. Mayou, Psychological consequences of road traffic accidents in children. Eur Child Adolesc Psychiatry, 1998. 7(2): p. 61-8.
43. Mayou, R.A., A. Ehlers, and B. Bryant, Posttraumatic stress disorder after motor vehicle accidents: 3-year follow-up of a prospective longitudinal study. Behav Res Ther, 2002. 40(6): p. 665-75.
44. O'Donnell, M.L., et al., Psychiatric morbidity following injury. Am J Psychiatry, 2004. 161(3): p. 507-14.
45. van Reekum, R., et al., Psychiatric disorders after traumatic brain injury. Brain Inj, 1996. 10(5): p. 319-27.
46. Mogotsi, M., D. Kaminer, and D.J. Stein, Quality of life in the anxiety disorders. Harv Rev Psychiatry, 2000. 8(6): p. 273-82.
47. Saarni, S.I., et al., Impact of psychiatric disorders on health-related quality of life: general population survey. Br J Psychiatry, 2007. 190: p. 326-32.
48. Polinder, S., et al., International variation in clinical injury incidence: exploring the performance of indicators based on health care, anatomical and outcome criteria. Accid Anal Prev, 2008. 40(1): p. 182-91.

Part I

Disability weights and health-related quality of life

2

Novel empirical disability weights to assess the burden of non-fatal injury

Juanita A. Haagsma, Ed F. Van Beeck, Suzanne Polinder, Nancy Hoeymans,
Saakje Mulder, Gouke J. Bonsel
Injury Prevention. 2008; 14(1): 5-10.

Abstract

Background Disability weights are necessary to estimate the disability component (years lived with disability, YLDs) of Disability Adjusted Life years. The original global burden of disease approach to deriving disability weights ignores temporary consequences of injury.
Objectives To develop and apply novel empirical disability weights to improve estimates of the non-fatal burden of injury.
Methods A set of 45 disability weights was derived for both permanent and temporary consequences of injury, using the annual profile approach. A population panel (n = 143) provided the values. The novel set of disability weights was then linked to epidemiological surveillance data on the incidence of injury in the Netherlands to calculate YLD resulting from permanent and temporary consequences of injury.
Results The empirical disability weights for injury consequences varied from minor (corneal abrasion, 0.004) to very severe (quadriplegia, 0.719) health loss. Increasing disability weights by level of severity were found, as illustrated by concussion (0.02), versus moderate brain injury (0.193), versus severe brain injury (0.540). Application of these new disability weights showed a 36% increase in YLD as the result of unintentional injury.
Conclusions YLD calculations based on global burden of disease disability weights underestimate the size of the injury problem by ignoring temporary health consequences. Application of our novel empirical disability weights, derived using the annual profile approach, may improve calculations on the burden of non-fatal injury.

Introduction

Worldwide, injuries are recognized as a major concern in clinical and public health. To reduce the number of injuries, a range of preventive interventions have been implemented, and more are being considered depending on opportunities and resources. However, resources available for prevention are limited. For policy makers, quantification of the burden of disease or injury and the subsequent determination of the relative attributes of avoidable risk factors to this burden are vital tools for priority setting purposes.

The burden of disease concept was introduced by the World Health Organization in 1996 [1]. It represents the total amount of health loss at the population level, and is quantified in terms of disability adjusted life years (DALYs). DALYs summarize (a) time lost through premature mortality (years of life lost), and (b) "healthy" time lost while living with a disability (years lived with disability (YLD)) in a single quantity. An essential factor for establishing YLDs is the disability weight, a value assigned to living with disability. This value, anchored between 0 and 1, reflects the impact of a specific health condition and is usually based on the preferences of a panel of judges. To this, dedicated preference measurement methods are used [2, 3].

For the global burden of disease (GBD) approach – which aims to estimate the total burden of disease worldwide as the sum of the burden of all separate diseases – a large set of disability weights had to be derived. Over 100 specific disease categories with a disability component have been distinguished. The initial approach used by Murray et al. was based on choices that have subsequently been challenged [4, 5]. Firstly, to arrive at the GBD disability weights, the disability-causing health states were presented to the panel of judges as a limited diagnostic description of the disease, derived from the terminology of the International Classification of Diseases (ICD). Apart from the disadvantage that only physicians could be involved in the evaluation, later studies showed that, even for expert panels, adding a generic description of functional health status to the diagnostic disease label is necessary to standardize the stimulus [6-8]. Secondly, in computing the DALYs, the GBD approach assumes independence between duration and disability and requires that the health state remains constant over time [9, 10], an assumption that is untenable for many diseases and injury consequences.

To overcome the operational limitations of the GBD-approach, Stouthard et al. developed an alternative approach, the annual profile approach (APA) [6]. The most important feature of the APA is that the course of the health state – the disability profile – is described over a period of 1 year time, alleviating the time constraints of the GBD-approach [10]. This allows derivation of disability weights for health states with an acute onset, episodic diseases such as epilepsy, and health states characterized by complex and heterogeneous recovery patterns, such as injury consequences.

Injuries encompass a wide range of health states that may have both temporary and permanent health consequences. Although temporary consequences of injury have a relatively small effect on health-related quality of an individual life, they are highly

prevalent, and, collectively their impact on health burden is significant [11-13]. By ignoring temporary consequences of injury, the GBD approach may typically underestimate the burden of injury and specific external causes, and therefore affect prioritization of resources for injury prevention. However, the GBD disability weights are currently used internationally to assess the burden of non-fatal injury, as alternative disability weights are not available [14].
In this study, our objective was to develop and apply novel empirical disability weights to improve estimates of the non-fatal burden of non-fatal injury.

Methods

Derivation of disability weights

Injury selection and health-state description

For the selection of injury consequences, the previously derived European injury classification EUROCOST was used [15]. This classification is compatible with the International Statistical Classification of Diseases, Injuries and Causes of Death – Ninth revision (ICD-9) classification system. The 39 EUROCOST injury groups were subdivided into 45 injury stages that are homogeneous in terms of healthcare use, disability, as well as treatment and prognosis.
Each of these 45 injury stages was presented on vignette (a preformatted A4 size sheet of paper that describes the functional consequences of an injury stage with four quadrants of disease information.) The left upper quadrant provided the disease label and clinical description. The upper right quadrant provided a generic description. For this generic description, we used the extended version of the EuroQol 5D (EQ-5D) [6, 16, 17]. The lower right quadrant described the course of the condition over 1 year time. Conditions with short duration were presented as a patient who, in an otherwise healthy year, experiences, for instance, the health effect of a concussion for 2 weeks; whereas conditions with long-term consequences were presented as a patient who experiences, for instance, the health effect of a moderate brain injury throughout the whole year. The lower left quadrant gave a visual representation of the injury.
For the generic descriptions of the 45 injury stages, we used data from a survey administered to a representative sample of patients who attended an emergency department in The Netherlands because of injury [18].

Health state valuations

To elicit preferences for the 45 injury stages, two valuation techniques were used: the visual analogue scale (VAS) and the time trade-off (TTO) [19] . The VAS valuation technique requires participants to score the injury stage on a vertical scale graded from 0 (worst imaginable health state) to 100 (best imaginable health state). With TTO, the

participants were asked how much time they would be willing to "trade" in order to be restored from the presented disease stage to full health. All vignettes were valued independently according to both methods.

Panel participants and data collection

The panel participants were lay people who applied to participate in the Mild Diseases and Ailments Study, conducted in 2003 [20]. For that study, 60 people were recruited from the general public through an advertisement in a newspaper that is freely available throughout The Netherlands. These persons were used as sampling framework for our study on injury disability weights. For this purpose, a random sample of 250 people was drawn and contacted by mail; 170 of these were willing to participate.
Initially, the panel participants attended a 3 h panel meeting, during which they evaluated 10 vignettes. Then they received a postal questionnaire. We developed three versions of the questionnaire, each containing 22 different vignettes in randomized order. The participants received €50 for participating in the study.

Data analysis

Firstly, we calculated VAS weights and TTO weights for each response using the following formulas:

VAS weight = 1 – (VAS score/100).

TTO weight = TTO score/365.

To establish whether the injury stages were ranked in a consistent order, the ranking of a vignette according to VAS and TTO weights was compared using the Spearman's rank correlation coefficient and the Pearson's rank correlation coefficient. We also performed regression analysis to determine whether age, sex and education level had independent significant effects on the VAS and TTO weights.
The disability weights were also calculated on the basis of both mean VAS and mean TTO weights using a log-transformation procedure [21]. Previous studies have shown that the valuation technique used to elicit preferences may affect the value of the preferences [21-23]. The transformation procedure aims to adjust for this potential effect of the valuation technique.

Calculation of burden of non-fatal injury (YLD)

The following surveillance systems provided national data on injury incidence and their consequences:

I. The Dutch Injury Surveillance system: a permanent registry of injuries treated at the Emergency Department of 17 hospitals [18].

II. The Second Dutch National Survey of General Practice: a nationwide survey, including data on visits to a general practitioner for injury; performed by the Netherlands Institute for Health Services Research (NIVEL) in 2001 [24].
III. A follow-up survey administered to a representative sample of patients wit injury up to2 years after treatment at the emergency department of a hospital [18].

To calculate the YLD, incidence data from the surveillance systems was multiplied by mean duration and the log-transformed disability weight derived from the panel study. Injury consequences for which at least 50% of the participants were not willing to "trade off" any time were considered to be trivial and were therefore excluded from the YLD calculation – that is, corneal abrasion and superficial injury, for which 72% and 75%, respectively, of the participants were not willing to trade off any time.

Results

Disability weights

A total of 143 participants attended the panel meeting, 136 (95%) of whom responded to the questionnaire. The response rates of the three versions of the postal questionnaire were 94%, 98% and 96%. Table 2.1 gives some details of the participants.

Table 2.1 Basic characteristics of the 143 participants

Characteristic	Mean (range) or n (%)
Age (years)	47.5 (19-82)
Sex	
Male	59 (41)
Female	84 (59)
Education level	
Low	25 (17)
Middle	61 (43)
High	55 (38)
Not available	2 (0.01)

Table 2.2 presents the mean and median VAS and TTO weights for the 45 injury stages. The results show that the participants evaluated the injury stages logically; lowest VAS and TTO weights were attributed to mild injury, such as superficial injury (mean VAS 0.09, mean TTO 0.01), and highest VAS and TTO weights to severe injuries, such as quadriplegia (mean VAS 0.89, mean TTO 0.64). This is illustrated by increasing disability weights by level of severity within specific types of injury. For example, concussion (mean VAS 0.20, mean TTO 0.01) is rated lower than moderate brain injury (mean VAS 0.55, mean TTO 0.27) and severe brain injury (mean VAS 0.80, mean TTO 0.33). Furthermore,

the results from table 2.2 show that, if injuries were subdivided into an acute phase and a chronic phase, the acute phase had higher VAS and TTO values (less preferable). For example, the acute phase of multiple injury without brain injury (mean VAS 0.65, mean TTO 0.27) is rated higher than the stable phase of multiple injury without brain injury (mean VAS 0.49, mean TTO 0.23). In contrast with the VAS, the TTO ranking of some of the severe injury stages did not correspond to the severity of the injury stage.

The Spearman correlation coefficient between mean VAS and TTO values was 0.97, indicating a similar rank ordering of the injury stages. Pearson's correlation coefficient was 0.93, indicating similar values of the injury stages.

The standard deviation of the TTO values (data not shown) was higher in the middle range, which was expected given the fixed end points of the scale. No significant effects of gender, age, or educational level on the VAS and TTO values of the injury stages were shown. Table 2.2 also shows the log-transformed disability weights.

Table 2.2 Visual analogue scale (VAS) and time trade-off (TTO) values and log-transformed disability weights (DW) for the 45 injury stages

Injury stage		VAS		TTO		
	n	mean	median	mean	median	DW[1]
Head injury						
Concussion	142	0.20	0.20	0.01	0.01	0.020
Moderate brain injury	43	0.55	0.55	0.27	0.15	0.193
Severe brain injury, acute	46	0.80	0.85	0.33	0.25	0.540
Severe brain injury, stable	44	0.74	0.75	0.35	0.29	0.429
Corneal abrasion				<0.0		
	44	0.07	0.05	1	0	0.004
Fracture of nose	43	0.13	0.10	0.01	<0.01	0.009
Fracture of jaw	46	0.27	0.26	0.03	0.02	0.038
Back injury						
Fracture of vertebrae	43	0.54	0.53	0.21	0.13	0.186
Back sprain	46	0.27	0.25	0.04	0.01	0.039
Whiplash	44	0.33	0.32	0.07	0.04	0.056
Paraplegia, acute	142	0.82	0.82	0.50	0.44	0.563
Paraplegia, stable	43	0.86	0.86	0.63	0.54	0.656
Quadriplegia, acute	46	0.89	0.90	0.51	0.50	0.713
Quadriplegia, stable	44	0.89	0.90	0.64	0.75	0.719
Injury of thorax						
Fracture of rib	43	0.29	0.27	0.04	0.04	0.045
Injury of upper extremity						
Fracture of clavicle	142	0.28	0.28	0.05	0.03	0.041

Fracture of upper arm	44	0.27	0.30	0.06	0.04	0.039
Fracture of forearm	47	0.25	0.25	0.03	0.05	0,062
Fracture of wrist	43	0.30	0.29	0.05	0.03	0.049
Fracture of finger	142	0.16	0.15	0.03	0.01	0.014
Dislocation of shoulder	46	0.29	0.27	0.04	0.02	0.043
Sprain of wrist	45	0.23	0.24	0.02	0.01	0.026
Traumatic amputation finger	46	0.41	0.45	0.10	0.05	0.048
Traumatic amputation thumb	43	0.47	0.50	0.20	0.12	0.135
Injury of pelvis						
Fracture of pelvis	43	0.50	0.51	0.15	0.12	0.155
Injury of lower extremity						
Fracture of hip	142	0.46	0.45	0.11	0.08	0.124
Fracture of lower leg	46	0.34	0.32	0.05	0.02	0.063
Fracture of ankle	46	0.34	0.32	0.03	0.02	0.061
Fracture of toe	43	0.18	0.19	0.01	0.01	0.017
Sprain of ankle	43	0.19	0.23	0.03	0.01	0.018
Dislocation of hip	44	0.39	0.40	0.07	0.05	0.083
Traumatic amputation toe	46	0.44	0.44	0.07	0.04	0.111
External injury						
Superficial injury	142	0.09	0.06	0.01	0	0.005
Open wound	46	0.14	0.10	0.01	<0.01	0.011
Small burn	44	0.12	0.10	0.01	0.01	0.008
Large burn, acute	43	0.69	0.71	0.42	0.36	0.357
Large burn, stable	44	0.60	0.60	0.46	0.46	0.248
Large burn, incl. face, acute	46	0.73	0.75	0.39	0.13	0.420
Large burn, incl. face, stable	43	0.77	0.77	0.51	0.46	0.479
Polytrauma						
Multiple injury, excl. brain, acute	46	0.65	0.69	0.27	0.15	0.304
Multiple injury, excl. brain, stable	44	0.49	0.45	0.23	0.12	0.145
Multiple injury, incl. brain, acute	43	0.78	0.80	0.47	0.40	0.487
Multiple injury, incl. brain, stable	45	0.76	0.80	0.29	0.23	0.461

[1] 0 = full health, 1 = worst possible health state

Calculation of burden of injury

The data sources showed that, in 2003 in The Netherlands, medical advice for the consequences of unintentional injuries was sought 1 714 000 times – that is, 11% of the population (population at risk 16,188,390). Almost one half (48%, 830 000) of the people

that sought medical advice were treated at the emergency department of a hospital. Of these, 5% (37 000) had long-term consequences (2 years after the injury).
There were 116 000 YLDs due to unintentional injuries, 52% of which were due to home and leisure injuries, 27% to traffic injuries, 16% to sport injuries, and 5% to occupational injuries.
Short-term consequences of injury contributed 31 000 YLDs to the total estimate of YLDs due to unintentional injuries, an increase of 36% on top of the 86 000 YLDs due to long-term consequences. For home and leisure injuries, traffic injuries, occupational injuries and sport injuries, short-term consequences contributed 17 000 YLDs (39%), 5500 YLDs (21%), 1900 YLDs (49%) and 6200 YLDs (50%), respectively (see table 2.3).

Table 2.3 Number of YLDs lost due to unintentional injury, by injury category

Injury category	Short-term consequences		Long-term consequences		Total
	n	YLD	n	YLD	YLD
Home and leisure	852,000	17,000	21,000	44,000	60,000
Traffic	235,000	5,500	5,800	26,000	31,000
Occupational	163,000	1,900	1,500	3,900	5,800
Sports	464,000	6,200	8,600	13,000	19,000
Total	1,714,000	31,000	37,000	86,000	116,000

Discussion

This study derived a set of disability weights for injury consequences, varying from minor to very severe health loss, using the APA. The empirically derived disability weights were used to calculate the burden of injury due to unintentional injuries in The Netherlands. We found that inclusion of injuries with temporary consequences increased the non-fatal burden of injury in The Netherlands by 36%. Short-term consequences were particularly important in sport and occupational injuries: 50% increase in the number of YLDs.
In this study, we used the relatively new approach, APA, rather than the standard QALY/DALY approach (SQA) applied in the original GBD study [1]. The SQA consists of adding up separately valued periods lived in a certain state of health. During these separate periods, state of health is assumed to remain constant. The implication of this approach is that health outcome characterized by dynamic and complex patterns such as injury consequences have to be separated into various periods of constant health.
In contrast with the SQA, the APA avoids the break down of health outcome into time periods because it describes the health outcome over the course of 1 year [10]. This 1-year perspective allows the derivation of disability weights for health outcome with short-term consequences, which have been ignored in previous burden estimates. To obtain a societal perspective, in this study, preferences for injury stages were obtained by a panel

derived from the general population., who evaluated each of the injury stages with both a rating scale and a trade-off technique. In contrast, the GBD-disability weights were obtained from an expert panel with a rating scale only [8].
The APA has been criticized by Vos, who stated that the disability weights derived with the APA overvalue diseases and injuries with a mild and rapid course [25]. According to this criticism, the burden of short-term consequences calculated in this study could be an overestimation, and the observed increase in burden of non-fatal injury may be an effect of the APA. However, Essink-Bot & Bonsel pointed out that the relatively high values of the disability weights derived with the APA are due to sensitivity problems at the end points of the scale [10]. Hence, the relatively high disability weights are a result of the lack of discrimination between conditions of low severity rather than the time presentation.
To avoid the aforementioned problem of disability weights of low-severity health states, it is important to distinguish trivial conditions from relevant conditions. Trivial conditions have a minor impact on health-related quality of individual lives, but may be highly prevalent. Inclusion of these conditions in the calculation of burden of disease or injury may result in a large number of YLDs, which stand out compared to severe, but much less common, conditions. As a result, trivial injuries may become the focus of attention rather than severe injuries. This problem was also pointed out by Sim and Mackie, who emphasized the importance of reliable indicators for measuring non-fatal injuries [26]. Therefore, in the current study we used the preference-based relevance threshold developed by Bonsel et al. to distinguish trivial injury consequences from minimal relevant injury consequences [20]. The preference-based threshold was met if the median TTO value is greater than 0 – that is, if at least 50% of the participants were willing to trade-off any time. If not, the health state was regarded as trivial according to the majority of the population and therefore excluded from the calculation of burden of injury. This threshold is based on societal preferences, which concurs with the societal perspective of the burden of disease concept, and it corresponds to the majority rules principle of most democratic voting systems. The results show that two injury types did not meet the threshold, namely corneal abrasion and superficial injury. These two injury types were therefore excluded from the YLD calculations. Inclusion of these two injury types would result in an increase of the number of YLD with an extra 3% (from 116 000 YLD to 120 000 YLD lost).
To meet the societal perspective of the burden of disease concept, it is important that the disability weights reflect preferences of the society under study [27]. The disability weights derived in the current study are based on preferences of a sample of the Dutch population. This, however, does not mean that the use of these disability weights is limited to The Netherlands. Two previous studies that compared health state valuations among residents of several countries found that ranking of health states is similar across countries [28, 29]. Secondly, the disability weights that resulted from the 1997 Dutch Disability Weights Study have been applied in burden of disease studies of many other countries, like Australia, New Zealand, Serbia and Spain [30-33].

Considering the duration and the complexity of the valuation task, the response rates of the panel study were high, with 57% of invited people actually participating in the panel session and 95% of the participants returning the follow-up questionnaire. These high response rates may be explained by the relatively high incentive of €50, the careful design of the study materials to enhance information processing, and the use of reminders for the follow-up questionnaire. Two panel studies with similar study design and incentives yielded equally high response rates [20, 34].

A limitation of this study was that, although the panel overall ranked the injury stages consistently and logically, the TTO values showed some unexpected inconsistencies for severe health states. With the TTO technique, participants seemed less able to discriminate between severe injury stages, suggesting that minor inconsistencies might be caused by the conceptually difficult trade-off valuation technique. The VAS on the other hand is conceptually less difficult, and its values did not show inconsistencies. However, the VAS lacks the trade-off feature; it does not ask for sacrifice something valuable in order to assess the undesirability of the health state. Consequently the preferences elicited with the VAS gives less information about the relative desirability of a health state compared to other health states and are therefore regarded as less appropriate for eliciting preferences and, for this reason, less appropriate for basing disability weights on [21]. To adjust for the minor TTO inconsistencies, the disability weights are based on both VAS and TTO values.

We conclude that the disability weights that we have derived in this study may be used to overcome some of the limitations of the GBD disability weights. Furthermore, we conclude that highly prevalent injuries with temporary consequences should not be ignored in burden of injury estimates, provided that there is a threshold of triviality.

Implications for prevention

The burden of disease concept is used increasingly for priority setting in healthcare and prevention. However, calculations of burden of disease based on GBD disability weights ignore temporary health consequences, resulting in an underestimation of the size of the injury problem. This novel derived set of disability weights for injury consequences includes highly prevalent short-term consequences, improving burden of injury estimates. This allows policy makers to base their decisions regarding resource allocation on a more complete burden of injury.

References

1. Murray, C.J.L. and A.D. Lopez, The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. 1996, Cambridge: Harvard University Press.
2. Drummond, M., et al., Methods for the Economic Evaluation of Health Care Programmes. 1997, Oxford, U.K.: Oxford Medical Publications.

3. Brazier, J., M. Deverill, and C. Green, A review of the use of health status measures in economic evaluation. J Health Serv Res Policy, 1999. 4(3): p. 174-84.
4. Murray, C.J. and A.K. Acharya, Understanding DALYs (disability-adjusted life years). J Health Econ, 1997. 16(6): p. 703-30.
5. Murray, C.J., Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ, 1994. 72(3): p. 429-45.
6. Stouthard, M.E., M.L. Essink-Bot, and G.J. Bonsel, Disability weights for diseases. A modified protocol and results for a Western European region. European Journal of Public Health, 2000. 10: p. 24-30.
7. Schwarzinger, M., et al., Cross-national agreement on disability weights: the European Disability Weights Project. Popul Health Metr, 2003. 1(1): p. 9.
8. Anand, S. and K. Hanson, Disability-adjusted life years: a critical review. J Health Econ, 1997. 16(6): p. 685-702.
9. Gold, M.R., D. Stevenson, and D.G. Fryback, HALYS and QALYS and DALYS, Oh My: similarities and differences in summary measures of population Health. Annu Rev Public Health, 2002. 23: p. 115-34.
10. Essink-Bot, M.L. and G.J. Bonsel, How to derive disability weights?, in Summary measures of population health: concepts, ethics, measurement and applications., C.J.L. Murray, A.D. Lopez, and J.A. Salomon, Editors. 2002, World Health Organization: Geneva.
11. Meerding, W.J., et al., Distribution and determinants of health and work status in a comprehensive population of injury patients. J Trauma, 2004. 56(1): p. 150-61.
12. Polinder, S., et al., Prevalence and prognostic factors of disability after childhood injury. Pediatrics, 2005. 116(6): p. e810-7.
13. McClure, R.J. and R.M. Douglas, The public health impact of minor injury. Accid Anal Prev, 1996. 28(4): p. 443-51.
14. Polinder, S., et al., Assessing the burden of injury in six European countries. Bull World Health Organ, 2007. 85(1): p. 27-34.
15. Lyons, R.A., et al., Methodological issues in comparing injury incidence across countries. Int J Inj Contr Saf Promot, 2006. 13(2): p. 63-70.
16. Krabbe, P.F., et al., The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. J Clin Epidemiol, 1999. 52(4): p. 293-301.
17. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group Health Policy, 1990. 16(3): p. 199-208.
18. Polinder, S., et al., Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma, 2007. 62(1): p. 133-41.
19. Torrance, G.W., W.H. Thomas, and D.L. Sackett, A utility maximization model for evaluation of health care programs. Health Serv Res, 1972. 7(2): p. 118-33.
20. Bonsel, G.J., M.F. Janssen, and E. Birnie, Mild diseases & Ailments Study. Ziektelast-bepaling van 27 vermoedelijk 'lichte' ziekten ten behoeve van rationele beperking van het geneesmiddelenpakket. 2004, University of Amsterdam, Academic Medical Center: Amsterdam.
21. Krabbe, P.F., M.L. Essink-Bot, and G.J. Bonsel, The comparability and reliability of five health-state valuation methods. Soc Sci Med, 1997. 45(11): p. 1641-52.
22. Lin, M.R., et al., Rating scale, standard gamble, and time trade-off for people with traumatic spinal cord injuries. Phys Ther, 2006. 86(3): p. 337-44.
23. Badia, X., et al., Feasibility, validity and test-retest reliability of scaling methods for health states: the visual analogue scale and the time trade-off. Qual Life Res, 1999. 8(4): p. 303-10.
24. Westert, G.P., et al., Monitoring health inequalities through general practice: the Second Dutch National Survey of General Practice. Eur J Public Health, 2005. 15(1): p. 59-65.

25. Vos T. The case against annual profiles for the valuation of disability weights. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD, editors. Summary measures of population health: concepts, ethics, measurement and application.Geneva: World Health Organization; 2002.
26. Sim, F. and P. Mackie, It were to be wished the flaws were fewer...but the main thing is, does it hold good measure? Public Health, 2002. 116(5): p. 251.
27. Field, M.J. and M.R. Gold, Summarising population health: Directions for the development and application of population health metrics. . 1998, Washington D.C.: Institute of Medicine: National Academy Press.
28. Ustun, T.B., et al., Multiple-informant ranking of the disabling effects of different health conditions in 14 countries. WHO/NIH Joint Project CAR Study Group. Lancet, 1999. 354(9173): p. 111-5.
29. Essink-Bot, M.L., et al., Cross-national comparability of burden of disease estimates: the European Disability Weights Project. Bull World Health Organ, 2002. 80(8): p. 644-52.
30. Mathers, C.D., et al., The burden of disease and injury in Australia. Bull World Health Organ, 2001. 79(11): p. 1076-84.
31. Jankovic, S., et al., The burden of disease and injury in Serbia. Eur J Public Health, 2007. 17(1): p. 80-5.
32. Norman R., et al., Revised burden of disease estimates for the comparative risk factor assessment, South Africa 2000. 2006, South African Medical Research Council: Cape Town.
33. Cubo, E., et al., Burden of disease related to Parkinson's disease in Spain in the year 2000. Mov Disord, 2005. 20(11): p. 1481-7.
34. Janssen, M.F., et al., Comparing the standard EQ-5D three-level system with a five-level version. Value Health, 2008. 11(2): p. 275-84.

3

Alternative approaches to derive disability weights for injuries: do they make a difference?

Juanita A. Haagsma, Suzanne Polinder, Ed F. Van Beeck, Saakje Mulder, Gouke J. Bonsel
Quality of Life Research. 2009; 18(5): 657-65

Abstract

Background In burden of disease studies, several approaches are used to assess disability weights, a scaling factor necessary to compute years lived with disability (YLD). The aim of this study was to quantify disability weights for injury consequences with two competing approaches, (a) standard QALY/DALY model (SQM) which derives disability weights from patient survey data and (b) the annual profile model (APM) which derives weights for the same patient data valued by a panel.

Methods Disability weights were assessed using (a) EQ-5D data from a postal survey among 8,564 injury patients 2½, 5, and 9 months after attending the emergency department, and (b) preferences of 143 laymen elicited with the time trade-off method.

Results Compared with APM, SQM disability weights were consistently higher. YLD calculated with SQM disability weights was more than three times higher compared to YLD calculated with APM disability weights, for mild injuries with short duration, this increase was six fold.

Conclusions The APM seems the preferred method in burden of injury studies that include mild conditions with a rapid course, since the SQM approach might overestimate the impact of the latter. The APM, however, might underestimate the impact of injury consequences, especially in case of severe injuries.

Introduction

Since the application of the concept in 1993, the disability adjusted life year (DALY) is used increasingly for priority setting in health care and prevention [1]. The DALY is a health gap measure that aggregates mortality and morbidity data, thus allowing comparison of population health status between countries as well as comparison between diseases within a country [2]. To aggregate mortality and morbidity data, years of life lost due to premature mortality (YLL) and years lived with disability (YLD) have to be established. Essential for the latter is the disability weight; a scaling factor that expresses the impact of a disease with a value ranging from 0, indicating best possible health state, through 1, indicating worst possible health state [3]. By multiplying the disability weight of a condition by its incidence and its average duration (or prevalence in case of chronic disease), the healthy time lost due to living with disability (YLD) is calculated.

Regarding disability weights there are two dominant approaches, both have been used in burden of disease studies. One of these approaches is to adopt existing disability weights from the Global Burden of Disease (GBD) study [4]. In order to compute YLD, the GBD disability weights are then applied using the standard QALY/DALY model (SQM). The SQM assumes independence between duration and disability and it requires that the health state remains fixed over time [5-7]. For health states with dynamic and/or complex patterns this assumption is untenable, since these health states in fact have to be separated into numerous parts.

A field abundant of dynamic recovery patterns with a wide variation in duration is the field of injuries. Moreover, the existing set of disability weights as published by Murray et al. lack a number of highly incident non-ignorable long-term injury consequences, which ultimately result in an underestimation of the total burden of injury [4]. To address both the issue of complexity over time and the issue of incomplete coverage of long-term sequelae, existing methods have to be adapted or extended.

At this stage, the SQM approach has been administered to several burden of injury studies [8-11]. This approach uses a two-step procedure to assign disability weights to health outcomes. Firstly, patients report their own health state using one of the available generic health state classification systems. These classification systems render the health state of an individual by the function level he/she reports on each of the domains. The weight of that health state is computed by a formula that firstly yields a partial weight score for each domain depending on the reported level, and subsequently adds partial weights which by definition fit within the 0.0-1.0 range. The partial weights of the formula, the tariff weights, are derived at an earlier stage from preference data of the population [12]. Commonly used systems are the EuroQol-5D (EQ-5D), the Health Utility Index, and the Quality of Well-Being Scale [13-15].

An alternative method used in burden of disease and injury studies is to obtain disability weights directly using the annual profile model (APM) [16, 17]. Unlike the SQM, the APM describes the course of the condition over 1-year time, allowing assessment of disability weights for health states characterized by an acute onset and complex patterns

of recovery. The measurement techniques to elicit the preference weights are, however, identical [16].
The aims of this study were to quantify differences between (a) disability weights for injury consequences derived from patient reported EQ-5D classification data using the SQM and (b) disability weights derived with panel elicitation using the APM, and subsequently compare YLD estimations calculated with both sets of disability weights given similar incidence data. Based on this comparison we aimed to decide on the preference of APM above SQM in injuries.

Methods

For the selection of the injury consequences, the EUROCOST injury classification was used [18]. This classification is compatible to the International Statistical Classification of Disease, Injuries and Causes of Death – ninth and tenth revision (ICD-9 and 10) classification system.
For assessment of the SQM and APM disability weights, data from two studies were used (1) the EQ-5D that was included in a questionnaire held among injury patients who attended the emergency department (ED) [9] and (2) health state valuations derived with APM panel elicitation using the visual analogue scale (VAS) and time trade-off (TTO) technique [17].

SQM disability weights

A sample of 8,564 injury patients aged 15 years and older was sent a postal questionnaire 2½ months after they attended the ED of a hospital in The Netherlands [9]. At 5 and 9 months, a follow-up questionnaire was sent to patients that responded to the preceding questionnaire. The sustained injuries varied from minor to severe and the sample included hospitalized and non-hospitalized patients. The sample of patients was stratified, over sampling patients with severe injuries. To measure functional outcome after injury, the questionnaires included the EQ-5D and the VAS. With the EQ-5D classification system, subjects describe their health in three levels of severity in the dimensions mobility, self-care, usual activities, pain/discomfort and anxiety/depression [19]. Subsequently, the reported health states were converted into utility scores using a pre-existing set of preference weights based on preference data from the general population of The Netherlands [20]. To adjust the data for non-response, a non-response analysis was performed by multivariate logistic regression, testing variables age, sex, type of injury, external cause of the injury, hospitalization and length of stay, health status and ambulance transport as possible determinants of non-response. The significant variables ($p<0.05$) were used to yield non-response weights for the sample of patients treated at the ED. The data were also adjusted for stratification of the sample of ED patients [9]. The resulting weighted data can be considered representative for an ED population of injury

patients in The Netherlands. We used the weighted data to calculate the EQ-5D utility scores and included only patients who filled out all three questionnaires. For the subsequent calculation of the EQ-5D disability weights, we used the population health index of the population of the United Kingdom, adjusted for age and sex [21]. In order to determine the SQM disability weights per injury group over 1-year time ($DW_{SQM,\ 1\ year}$), the resulting 2½-month (DW_{T1}), 5-month (DW_{T2}) and 9-month (DW_{T3}) disability weights were aggregated with the following formula:

$DW_{SQM,\ 1\ year} = (DW_{T1} + DW_{T2} + DW_{T3}) / 3.$

This formula was also used to calculate the aggregated VAS values per injury group over 1-year time.

APM disability weights

The APM disability weights were derived using the Dutch Disability Weights protocol with two important modifications [22]. Firstly, a population panel rather than a panel of medical experts provided the values. The population panel (n = 143) was randomly selected from an existing panel of 560 people that was recruited from the general public through an advertisement in a free newspaper that is available throughout the Netherlands. Secondly, the VAS and the TTO preference measurement methods were used to value a number of injury related health states. The VAS valuation technique requires participants to score the disease stage on a vertical thermometer graded from 0 (worst possible health state) to 100 (worst possible health state). The TTO asks participants how many days of 1 year in full health, they are willing to trade in order to be restored from the presented disease stage. Similarly to the Dutch Disability Weights study, the health states were described on a standardized 210 by 297 mm (A4) sized vignette which contained disease specific information in laymen terminology, a generic EQ-5D profile of the health state and an annual profile. This annual profile describes the course of the health state – the disability profile – over 1-year time, allowing assessment of diseases and injuries with rapid course and/or complex recovery patterns [16]. Additionally, information on the location of the injury and physical alterations caused by the health state was provided. To enhance information processing, we used graphics and colours in the description and we intensively explained the health state descriptions during the panel session. Each participant valued 32 health states: 10 health states during a 3 h panel session and subsequently 22 health states in a questionnaire, they received at home. The order of the presented health states was randomized.

Analysis

The questionnaire rendered SQM disability weights for 32 injury groups. For 11 of these injury groups, no matching APM disability weights were available. Hence, 21 injury groups were included in the comparison between SQM and APM disability weights. For each injury group absolute difference between SQM and APM disability weights as well as

SQM/APM disability weight ratio was calculated. A regression analysis was performed to determine if age and sex had significant effects APM disability weights. The 21 injury groups were categorized into three severity classes. These injury severity classes were grouped post hoc according to the calculated APM disability weights as previously tested by an international expert group [23]. Three injury severity classes were distinguished: mild, moderate and severe, using 0.03 and 0.10 as cut-off points. Differences between SQM and APM and ratios were calculated by injury severity class. Per injury severity class, YLDs were computed with both SQM and APM disability weights to compare the proportion YLDs lost due to mild, moderate and severe injury consequences. To calculate YLDs lost in the first year after the injury, the incidence data (year 2002) derived from the Dutch Injury Surveillance system were used. The Dutch Injury Surveillance system is a permanent registry of injury victims treated at the ED of 15 hospitals in The Netherlands [18]. The Spearman's rank correlation coefficient was calculated to test if the SQM and APM ranking based on the mean disability weights of the 21 injury consequences were associated. The Pearson correlation coefficient was used to test whether the distributions of the SQM and APM disability weights assigned to the injury consequences were correlated.

Results

Respondents

Of the 1,392 injury patients that completed the patient surveys 2½, 5 and 9 months after attending the ED 53% was male and mean age was 43 years old. The panel study was attended by 143 lay persons. Of these lay persons 59 (41%) was male and mean age was 48 years old.

Comparison of SQM and APM disability weights

Table 3.1 shows that the patient reported 2½ month disability weights were highest and the 9 month disability weights were lowest for all injury groups except spinal cord injury. With both SQM and APM disability weights, 0 indicates best possible health state and 1 indicates worst possible health state.

The mean SQM and APM disability weights are presented in Table 3.2. Mean SQM disability weights ranged from 0.03 (eye injury and open wound) to 0.55 (spinal cord injury) and mean APM disability weights ranged from 0.002 (eye injury) to 0.57 (spinal cord injury). For 19 of the total 21 injury consequences, mean SQM disability weights were higher compared to APM disability weights; the difference ranging from 0.004 (fracture clavicula/ scapula) to 0.09 (dislocation/sprain/strain hip and fracture knee/lower leg), with a mean difference of 0.04. The mean SQM disability weights for eye injury and superficial injury were both thirteen times higher than the AMP disability

weights. For the two injury groups concussion and fracture of hand/fingers, the SQM disability weights were more than five times higher.

Table 3.1 Disability weights derived with the standard QALY/DALY model (SQM[1]) at 2½ (T1), 5 (T2) and 9 (T3) month follow-up, per injury group

	T1		T2		T3	
Injury group	DW[2]	SD[3]	DW[2]	SD[3]	DW[2]	SD[3]
Head injury						
Concussion	0.14	0.22	0.08	0.19	0.04	0.11
Eye injury	0.02	0.06	0.05	0.15	0.01	0.03
Fracture facial bones	0.06	0.09	0.06	0.12	0.01	0.03
Injury of spinal cord/thorax/pelvis						
Fract/disl/spr/str vertebral column	0.21	0.21	0.13	0.18	0.10	0.14
Spinal cord injury	0.53	0.23	0.58	0.28	0.54	0.45
Fracture rib/sternum	0.11	0.20	0.09	0.16	0.08	0.18
Fracture pelvis	0.19	0.20	0.14	0.18	0.16	0.15
Injury of upper extremity						
Fracture clavicula/scapula	0.11	0.14	0.03	0.08	0.02	0.08
Fracture elbow/fore arm	0.15	0.16	0.02	0.06	0.03	0.07
Fracture wrist	0.09	0.14	0.05	0.13	<0.01	0.03
Fracture hand/fingers	0.12	0.15	0.07	0.16	0.07	0.21
Disl/sprain/strain shoulder/elbow	0.12	0.13	0.05	0.10	0.04	0.09
Disl/sprain/strain wrist/hand/fingers	0.09	0.19	0.05	0.16	0.03	0.11
Injury of lower extremity						
Fracture hip	0.29	0.30	0.22	0.26	0.18	0.28
Fracture knee/lower leg	0.24	0.21	0.11	0.17	0.07	0.14
Fracture ankle	0.16	0.18	0.08	0.13	0.06	0.11
Fracture foot/toes	0.06	0.11	0.04	0.09	0.03	0.07
Disl/sprain/strain ankle/foot	0.12	0.13	0.07	0.13	0.09	0.16
Disl/sprain/strain hip	0.22	0.23	0.13	0.21	0.15	0.26
External injury						
Superficial injury	0.12	0.19	0.07	0.16	0.05	0.15
Open wound	0.07	0.11	0.02	0.07	0.01	0.02

[1] SQM = disability weight derived from patient-reported EQ-5D data, adjusted for age and sex of the patient.

[2] DW = disability weights; 0 = full health, 1 = worst possible health state

[3] SD = standard deviation

Table 3.2 Disability weights derived with the standard QALY/DALY model (SQM) and the annual profile model (APM), per injury group

	SQM		APM			
Injury group	DW[1]	SD[2]	DW[1]	SD[2]	Δ[3]	Ratio
Head injury						
Concussion	0.088	0.13	0.015	0.02	0.07	5.9
Eye injury	0.027	0.06	0.002	0.01	0.02	13.4
Fracture facial bones	0.041	0.06	0.018	0.04	0.02	2.3
Injury of spinal cord/thorax/pelvis						
Fract/disl/spr/str vertebral column	0.147	0.14	0.108	0.17	0.04	1.4
Spinal cord injury	0.551	0.29	0.567	0.32	0.02	1.0
Fracture rib/sternum	0.092	0.17	0.045	0.04	0.05	2.0
Fracture pelvis	0.155	0.15	0.150	0.13	0.01	1.0
Injury of upper extremity						
Fracture clavicula/scapula	0.054	0.09	0.050	0.07	<0.01	1.1
Fracture elbow/fore arm	0.065	0.07	0.031	0.06	0.03	2.1
Fracture wrist	0.047	0.07	0.054	0.11	0.01	0.9
Fracture hand/fingers	0.086	0.16	0.016	0.05	0.07	5.4
Disl/sprain/strain shoulder/elbow	0.073	0.08	0.036	0.06	0.04	2.0
Disl/sprain/strain wrist/hand/fingers	0.057	0.15	0.027	0.05	0.03	2.1
Injury of lower extremity						
Fracture hip	0.231	0.27	0.202	0.17	0.03	1.1
Fracture knee/lower leg	0.139	0.15	0.049	0.09	0.09	2.8
Fracture ankle	0.102	0.11	0.056	0.08	0.05	1.8
Fracture foot/toes	0.044	0.08	0.014	0.02	0.03	3.1
Disl/sprain/strain ankle/foot	0.093	0.13	0.026	0.03	0.07	3.6
Disl/sprain/strain hip	0.166	0.19	0.072	0.08	0.09	2.3
External injury						
Superficial injury	0.079	0.15	0.006	0.01	0.07	13.2
Open wound	0.032	0.06	0.013	0.01	0.02	2.5

[1] DW = disability weight; 0 = full health, 1 = worst possible health state

[2] SD = standard deviation

[3] Δ = absolute difference between SQM and APM disability weights

Table 3.3 shows that the largest absolute differences and SQM/APM disability weights ratio were found for mild injuries (mean difference of 0.05, mean ratio of 4). Correlation coefficients between SQM and APM disability weights were high, Pearson's correlation coefficient was 0.93 and Spearman correlation coefficient was 0.65. No significant effects of age and sex on the TTO values were shown.

Table 3.3 Disability weights derived with the standard QALY/DALY model (SQM) and the annual profile model (APM), per severity class

Injury severity class	SQM	APM	Δ[1]	ratio
Mild (APM disability weight < 0.03)	0.061	0.015	0.05	4.0
Moderate (APM disability weight 0.03-0.10)	0.092	0.049	0.04	1.9
Severe (APM disability weight > 0.10)	0.271	0.257	0.01	1.1

[1] Δ = absolute difference between SQM and APM disability weights

Comparison of VAS values

In Table 3.4, the mean VAS values derived from the injury patients and the population panel are presented. The VAS values were lowest for eye injury (injury patients 0.12; population panel 0.07) and highest for spinal cord injury (injury patients 0.59; population panel 0.87). Except for injury groups eye injury, dislocation/sprain/strain ankle/foot and superficial injury, the patient reported VAS values were lower compared to the mean VAS values of the corresponding injury groups derived from the population panel, with a mean difference of 0.09. Pearson's correlation coefficient between mean SQM and APM VAS values was 0.90 and Spearman correlation coefficient was 0.75.

Comparison of YLD estimations

Table 3.5 shows the YLD estimations calculated with (a) the set of SQM disability weights and (b) the set of APM disability weights.
Application of the APM disability weights resulted in 16,947 YLDs, whereas application of the SQM disability weights resulted in 54,159 YLDs, an increase of 320% compared to the YLD estimation with APM disability weights.
With SQM disability weights, most YLDs were caused by to superficial injury (23,219 YLDs) and dislocation/sprain/strain of ankle and foot (4,543 YLDs). Together these two injury consequences accounted for 51% of the YLDs. This in contrast to the YLD estimation calculated with APM disability weights, where superficial injury (1,763 YLDs) and dislocation/sprain/strain of ankle and foot (1,270 YLDs) together accounted for only 18% of YLDs. With the APM disability weights, most YLDs were caused by hip fracture (3,140 YLDs), contributing 19% of YLDs.

Table 3.4 Mean SQM VAS values reported by injury patients and mean APM VAS values derived from the population panel, per injury group

	SQM		APM			
Injury group	VAS[1]	SD[2]	VAS[1]	SD[2]	Δ[3]	ratio
Head injury						
Concussion	0.17	0.16	0.20	0.12	0.03	0.9
Eye injury	0.12	0.09	0.07	0.11	0.05	1.7
Fracture facial bones	0.18	0.11	0.20	0.13	0.02	0.9
Injury of spinal cord/thorax/pelvis						
Fract/disl/spr/str vertebral column	0.33	0.17	0.38	0.17	0.05	0.9
Spinal cord injury	0.59	0.20	0.87	0.11	0.28	0.8
Fracture rib/sternum	0.22	0.14	0.29	0.13	0.07	0.8
Fracture pelvis	0.36	0.14	0.50	0.10	0.14	0.7
Injury of upper extremity						
Fracture clavicula/scapula	0.15	0.15	0.28	0.11	0.13	0.5
Fracture elbow/fore arm	0.14	0.14	0.25	0.09	0.11	0.6
Fracture wrist	0.12	0.12	0.30	0.12	0.18	0.4
Fracture hand/fingers	0.14	0.13	0.16	0.11	0.02	0.9
Disl/sprain/strain shoulder/elbow	0.18	0.15	0.29	0.11	0.11	0.6
Disl/sprain/strain wrist/hand/fingers	0.15	0.13	0.23	0.12	0.08	0.7
Injury of lower extremity						
Fracture hip	0.41	0.20	0.46	0.13	0.05	0.9
Fracture knee/lower leg	0.22	0.17	0.34	0.14	0.12	0.6
Fracture ankle	0.21	0.17	0.34	0.13	0.13	0.6
Fracture foot/toes	0.14	0.10	0.18	0.08	0.04	0.8
Disl/sprain/strain ankle/foot	0.21	0.18	0.19	0.10	0.02	1.1
Disl/sprain/strain hip	0.32	0.19	0.39	0.12	0.07	0.8
External injury						
Superficial injury	0.21	0.17	0.09	0.07	0.12	2.3
Open wound	0.12	0.13	0.14	0.13	0.02	0.9

[1] VAS = VAS value; 0 = full health, 1 = worst possible health state

[2] SD = standard deviation

[3] Δ = absolute difference between SQM and APM disability weights

Table 3.5 Incidence and YLD estimations calculated with standard QALY/DALY model (SQM) and annual profile model (APM) disability weights, per injury group

		SQM		APM	
Injury group	Incidence	YLD	%	YLD	%
Head injury					
Concussion	15,000	1,349	2.5	230	1.4
Eye injury	25,000	687	1.3	51	0.3
Fracture facial bones	7,400	303	0.6	133	0.8
Injury of spinal cord/thorax/pelvis					
Fract/disl/spr/str vertebral column	4,400	650	1.2	477	2.8
Spinal cord injury	460	256	0.5	264	1.6
Fracture rib/sternum	5,000	463	0.9	227	1.3
Fracture pelvis	3,200	499	0.9	483	2.8
Injury of upper extremity					
Fracture clavicula/scapula	16,000	869	1.6	804	4.7
Fracture elbow/fore arm	27,000	1,727	3.2	824	4.9
Fracture wrist	41,000	1,926	3.6	2,213	13.1
Fracture hand/fingers	45,000	3,881	7.2	722	4.3
Disl/sprain/strain shoulder/elbow	15,000	1,095	2.0	540	3.2
Disl/sprain/strain wrist/hand/fingers	12,000	704	1.3	334	2.0
Injury of lower extremity					
Fracture hip	16,000	3,591	6.6	3,140	18.5
Fracture knee/lower leg	13,000	1,857	3.4	654	3.9
Fracture ankle	18,000	1,814	3.3	996	5.9
Fracture foot/toes	29,000	1,261	2.3	401	2.4
Disl/sprain/strain ankle/foot	49,000	4,543	8.4	1,270	7.5
Disl/sprain/strain hip	3,000	492	0.9	213	1.3
External injury					
Superficial injury	290,000	23,219	42.9	1,763	10.4
Open wound	93,000	2,972	5.5	1,207	7.1
Total	730,000	54,159	100	16,947	100

As shown in Figure 3.1, with the application of SQM disability weights, the majority (72%, 38,920 YLDs) of the total number of YLDs were lost due to mild injuries with a rapid course. With the APM disability weights 36% (6,112 YLDs) of the total number of YLDs were lost due to mild injuries.

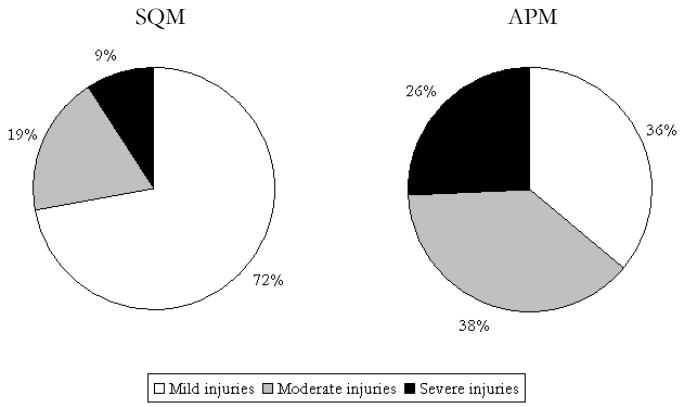

Figure 3.1 Percentage YLDs due to mild, moderate and severe injuries, calculated with the standard QALY/DALY model (SQM) and annual profile model (APM) disability weights, per injury group.

Discussion

The results showed that, although ranking of both sets of disability weights were concordant, the disability weights obtained with the SQM were consistently higher compared to disability weights obtained with the APM. The difference was relatively large for mild injuries with a rapid course, such as eye injury and superficial injury. Application of the SQM disability weights resulted in over three times as many YLDs lost in the first year after the injury compared to the YLD estimation with APM disability weights. For mild injury consequences this increase was six fold.

Unexpectedly, the VAS values showed the opposite compared to the disability weights obtained with the TTO. With VAS the population panel valued the health states worse than the patients did. Contrastingly to the TTO, the VAS valuation technique is not choice-based because it does not require the participants to make a trade-off between something valuable, time in case of the TTO, and health. However, this trade-off provides essential information about the relative (un)desirability of a certain health state compared to other health states. This makes the TTO values more appropriate for burden of disease studies.

The commonly used SQM method consists of adding up separately valued periods lived in a certain health state. It assumes that the value of the health state is not affected by the duration of the health state, nor by the sequence of the health states. Furthermore, during

the separate periods the health state is assumed to remain constant [5, 6]. To meet the assumption of constant disability, in order to assess the SQM disability weights we aggregated three periods of equal length valued at fixed points in time that were similar for each of the injury consequences. However, mild injuries with a rapid course consist of a sequence of much shorter periods of different disability levels. As a consequence, assuming constant health for a relatively long period might overestimate the actual impact of the injury consequences.

Reporting bias might also have contributed to the relatively high SQM disability weights. The SQM disability weights were derived from self-reported EQ-5D health status. Self-reported health status, however, might differ from the actual health status. Grootendorst et al. showed that respondents reported more dysfunction in self-completed questionnaires compared to interview-administered responses [24]. A factor that affects the responses to self-reported heath questions are preceding questions [25]. In the questionnaire that was sent to the injury patients, the EQ-5D was preceded by questions regarding the cause and immediate consequences of the sustained injury. This may have caused the injury patients to overemphasize their level of dysfunction at follow up.

A third aspect that may have affected the SQM disability weights is the baseline information used to calculate the disability weights. In the current study, we used population utility scores as a baseline. However, Cameron et al. showed that pre-existing morbidity in a cohort of injured patients is higher compared to non-injured individuals [26]. These findings of high pre-existing morbidity among injured patients are accorded by Wardle et al. and Polinder et al. [9, 27]. This implies that pre-injury utility scores are in fact lower than utility scores of the population. Therefore, using population utility scores as a baseline rather than pre-injury utility scores results in larger differences in health-related quality of life, and consequently higher SQM disability weights. Additionally, it should be noted that we have used UK population utility scores as a baseline, because EQ-5D population utility scores for The Netherlands are not available. The population health index of the UK, however, may not be comparable to the Dutch population health index.

For the calculation of the APM disability weights, a baseline utility is not required. Moreover, the APM avoids the assumption of constant disability. Rather than aggregating separately valued periods, the APM describes the disability profile of the condition – with generic as well as disease specific information – over the course of 1 year [16]. The APM allows a fixed preference-based threshold to distinguish trivial from minimal relevant disease [28]. As shown by the results, the relatively high values of the SQM disability weights for mild injuries with a rapid course in combination with high incidence results in a larger number of total YLDs, and may lead to policy priority of these mild injuries above severe, less frequently occurring injury consequences. This bias in the application of burden of disease estimates in prioritization issues is avoided by the APM threshold, which implies that only if more than 50% of participants are willing to trade off any time, injuries are regarded as relevant [29]. If the cut-off point is not met, the injuries are excluded from the burden of injury calculation. In the current study, two injury groups

did not meet the preference-based threshold, namely eye injury and superficial injury. If these two injury groups are excluded from the burden of disease calculations, the total number of YLDs calculated with the set of SQM disability weights will decrease from 54,159 to 30,253, with 49% of YLDs lost to mild injuries. The burden of disease calculated with APM disability weights will decrease from 16,947 to 15,133 YLDs, with 22% lost due to mild injuries.

On the other hand, the APM panel elicited disability weights have several limitations. Firstly, for any new health state, health state valuations have to be obtained by a new panel study in order to derive the disability weights. A second limitation of APM is that, although EQ-5D data from actual patients is used for the description of the health state, the APM disability weights are not able to capture the heterogeneity of the injury consequences as well as SQM disability weights, which are based on individual patient data. For instance, the consequences and duration of an open wound are highly dependent on the size of the wound and the location on the body. The APM health state description of open wound does not capture this variation. Moreover, several studies showed that variation in injury consequences increases with injury severity and duration [10, 30]. As a result, actual health states of injury patients may differ considerably from the health state descriptions valued with the APM, especially in case of severe injury consequences.

It should also be noted that although laypeople may be able to value highly incident injury consequences such as superficial injury and wrist fracture, it may be difficult for them to fully comprehend the effect of less frequently occurring injury consequences on the daily life of a patient living in the particular health state. Inconsistencies in the TTO valuations of severe injury consequences indicate that laypeople are less able to discriminate between severe injury consequences when using the conceptually difficult TTO valuation technique [17]. Hence, to calculate YLDs for severe long-term injury consequences, SQM disability weights might be preferable. For health outcome with dynamic or complex patterns, like mild injuries with short duration, the SQM seems to be less appropriate, since it results in relatively high disability weights that seem to overestimate the consequences.

The difference in VAS scores derived from injury patients and the population panel, with patients valuing their own health state as less severe compared, corresponds to the results found in a previous study [31]. A meta-analysis of 33 studies, however, showed that there were no significant differences between patient and non-patient preferences [32]; though it should be noted that the majority of the studies included in the meta-analysis concerned patients with chronic conditions, whereas the current study addressed injuries of mainly short duration. This disparity in duration of health consequences may have affected the patients' valuation of their own health state, since patients adapt to their health states. The effect of adaptation is especially found with chronically ill patients [33, 34]. The fact that the difference between SQM and APM disability weights is smaller for severe injuries with relatively long duration, like spinal cord injury, compared to mild injuries of short duration may also be due to adaptation to a certain health state.

Nevertheless, the present findings should be interpreted with caution because they are based on two separate datasets that did not allow a direct comparison of the data and because of the aspects mentioned above that may have affected the disability weights. The SQM disability weights might be considerably lower if a pre-injury baseline was used rather than a population baseline, and if the time interval to measure the health status of injury patients was more appropriate; although it is impossible to measure the health status of injury patients at the optimal time interval for the numerous consequences of injury. The values of the APM disability weights on the other hand might be higher if they were obtained from a panel of injury patients instead of a population panel, since injury patients have actually experienced the shock of accidentally sustaining an injury and the impact of its consequences on daily life.

We conclude that the approach used to assess disability weights does make a difference, and in their turn yield considerable differences in YLD calculations. The APM seems the preferred method in burden of injury studies that include mild conditions with a rapid course, since the SQM approach yields relatively high values that may overestimate the impact of the latter. The APM on the other hand may underestimate the impact of injury consequences, especially in case of severe injuries. Nonetheless, in comparing disease burden estimates between diseases or countries differences may be attributed to the methodological choice of disability weights applied in the DALY calculation rather than differences in incidence or prevalence.

References

1. Worldbank, World Development Report 1993: Investing in Health. 1993, New York: Oxford University Press.
2. Murray, C.J.L., A.D. Lopez, and C.D. Mathers, Summary measures of population health: concepts, ethics, measurement and applications. 2002, Geneva: World Health Organization.
3. Murray, C.J. and A.K. Acharya, Understanding DALYs (disability-adjusted life years). J Health Econ, 1997. 16(6): p. 703-30.
4. Murray, C.J.L. and A.D. Lopez, The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. 1996, Cambridge: Harvard University Press.
5. Pliskin, J.S., D.S. Shepard, and M. Weinstein, Utility functions for life years and health status. Operations Research, 1980. 28(1): p. 206-224.
6. Johannesson, M., B. Jonsson, and G. Karlsson, Outcome measurement in economic evaluation. Health Econ, 1996. 5(4): p. 279-96.
7. Tsuchiya, A. and P. Dolan, The QALY model and individual preferences for health states and health profiles over time: a systematic review of the literature. Med Decis Making, 2005. 25(4): p. 460-7.
8. Segui-Gomez, M. and E.J. MacKenzie, Measuring the public health impact of injuries. Epidemiol Rev, 2003. 25: p. 3-19.
9. Polinder, S., et al., Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma, 2007. 62(1): p. 133-41.

10. Holtslag, H.R., et al., Determinants of long-term functional consequences after major trauma. J Trauma, 2007. 62(4): p. 919-27.
11. Tilford, J.M., et al., Child health-related quality of life following neurocritical care for traumatic brain injury: an analysis of preference-weighted outcomes. Neurocrit Care, 2007. 7(1): p. 64-75.
12. Brazier, J., M. Deverill, and C. Green, A review of the use of health status measures in economic evaluation. J Health Serv Res Policy, 1999. 4(3): p. 174-84.
13. Kaplan, R.M., et al., Comparison of measured utility scores and imputed scores from the SF-36 in patients with rheumatoid arthritis. Med Care, 2005. 43(1): p. 79-87.
14. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group Health Policy, 1990. 16(3): p. 199-208.
15. Torrance, G.W., W.H. Thomas, and D.L. Sackett, A utility maximization model for evaluation of health care programs. Health Serv Res, 1972. 7(2): p. 118-33.
16. Essink-Bot, M.L. and G.J. Bonsel, How to derive disability weights?, in Summary measures of population health: concepts, ethics, measurement and applications., C.J.L. Murray, A.D. Lopez, and J.A. Salomon, Editors. 2002, World Health Organization: Geneva.
17. Haagsma, J.A., et al., Novel empirical disability weights to assess the burden of non-fatal injury. Injury Prevention., 2008. 14(5): p. 1-10.
18. Lyons, R.A., et al., Methodological issues in comparing injury incidence across countries. Int J Inj Contr Saf Promot, 2006. 13(2): p. 63-70.
19. Brooks, R., EuroQol: the current state of play. Health Policy, 1996. 37(1): p. 53-72.
20. Lamers, L.M., et al., The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ, 2006. 15(10): p. 1121-32.
21. Kind, P., G. Hardman, and S. Macran, UK population norms for EQ-5D. 1999, Centre of Health Economics, University of York: York.
22. Stouthard, M.E., M.L. Essink-Bot, and G.J. Bonsel, Disability weights for diseases. A modified protocol and results for a Western European region. European Journal of Public Health, 2000. 10: p. 24-30.
23. Polinder, S. and e. al., International variation in clinical injury incidence: exploring the performance of indicators based on health care, anatomical and outcome criteria. Accident Analysis and Prevention, In press.
24. Grootendorst, P., D. Feeny, and W. Furlong, Health Utilities Index Mark 3: evidence of construct validity for stroke and arthritis in a population health survey. Med Care, 2000. 38(3): p. 290-9.
25. Crossley, T.F. and S. Kennedy, The reliability of self-assessed health status. J Health Econ, 2002. 21(4): p. 643-58.
26. Gabbe, B.J., et al., Preinjury status: are orthopaedic trauma patients different than the general population? J Orthop Trauma, 2007. 21(4): p. 223-8.
27. Wardle, T., P. Driscoll, and M. Woolford, Medical problems in major trauma. Injury, 1995. 26: p. 137.
28. Bonsel, G.J., M.F. Janssen, and E. Birnie, Mild diseases and ailments study (MiDAS). Ziektelastbepaling van 27 vermoedelijk 'lichte' ziekten ten behoeve van rationele beperking van het geneesmiddelenpakket. 2003, Academic Medical Center, University of Amsterdam: Amsterdam.
29. Haagsma, J.A., et al., Disability Adjusted Life Years and minimal disease: application of a preference-based relevance criterion to rank enteric pathogens. Popul Health Metr, 2008. 6(1): p. 7.
30. Cornwall, R., et al., Functional outcomes and mortality vary among different types of hip fractures: a function of patient characteristics. Clin Orthop Relat Res, 2004(425): p. 64-71.
31. Boyd, N.F., et al., Whose utilities for decision analysis? Med Decis Making, 1990. 10(1): p. 58-67.

32. Dolders, M.G., et al., A meta-analysis demonstrates no significant differences between patient and population preferences. J Clin Epidemiol, 2006. 59(7): p. 653-64.
33. Adang, E.M., et al., Quality-of-life ratings in patients with chronic illnesses. Jama, 1997. 277(13): p. 1038.
34. Postulart, D. and E.M. Adang, Response shift and adaptation in chronically ill patients. Med Decis Making, 2000. 20(2): p. 186-93.

4

Comparison of generic and disease specific health state valuations by a laymen panel

Juanita A. Haagsma, M.F. Janssen, Erwin Birnie, Gouke J. Bonsel

Conference proceedings of the EuroQol Plenary Meeting 2005

Abstract

Background The objective of this study was to explore the effects of adding disease specific information to generic health state descriptions on health state valuations by a panel of lay people.

Methods Twenty-three different health states of common diseases were presented for valuation. Each of these health states consisted of a EQ-5D profile and were presented with and without a disease label and a formalized clinical description of the disease. A panel of lay people (n = 105), recruited from the general public, valued the health states with the Visual Analogue Scale (VAS).

Results For 19 health states, mean VAS values for health states with a disease specific information were numerically lower ('more severe') compared to those without this information ($p < 0,001$), with a mean difference of 0,10 (range: -0,08 to 0,28). The difference in VAS values was greater for less severe diseases. Standard deviations did not differ between health states with and without disease specific information, indicating similar precision.

Conclusions We found that the panel of lay people valued health states with disease specific information as more severe, especially for mild diseases. This indicates that disease specific health state descriptions contain information about prominent symptoms not reflected in the generic health state, consequently causing a valid shift.

Introduction

Information on the health status of a population is essential input for burden of disease studies and cost-effectiveness analysis of interventions. To quantify population's health status, summary measures of population health may be used, which summarize fatal and non-fatal health outcome into a single number [1]. Time lost due to premature mortality can be added to time lost due to morbidity using disability weights. The disability weight reflects the impact of a condition and its value is based on the preferences of a panel of judges stated towards a set of health states [2]. The preferences towards health states are anchored between 0, indicating indifference between the health state and full health, and 1, indicating indifference between the health state and worst possible health state. The precise number is commonly referred to as disability weight.

To derive disability weights, several choices have to be made considering the approach to quantify the preferences of the panel of judges [3]. One of these choices concerns the description of the health state, which may be described as (1) disease specific or (2) generic. Disease specific health state descriptions indicate the cause, the specific health effects and the treatment of the condition. A generic health description system describes functional health without regard to the underlying condition.

A widely used generic health state description system is the EuroQol classification system (EQ-5D), which describes health in the dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression [4, 5]. Generic health state descriptions, like EQ-5D health profiles, allow comparison across diseases and interventions. However, disease specific health state descriptions are more sensitive for the detection and quantification of small changes [6]. This indicates that disease specific health state descriptions provide information that is not reflected in the generic health states but which matters for health state valuation. This was also found by Stouthard et al. [7]. In their study, Stouthard et al. combined disease specific information with EQ-5D data [7]. They found that lower back pain and prostate cancer, health states with a similar EQ-5D profile and a different disease label, yielded different values. However, because medical experts provided the values, the disease label might have created a mental image of the typical patient, which the medical experts valued rather than the actual information that was presented to them.

Contrary to medical experts, lay people are unfamiliar with most conditions and its consequences. As a result, lay people may be less affected by the information that is provided by the disease label compared with medical experts. To our knowledge the effect of disease specific information on health state valuation of lay people has not yet been examined. The aim of this study was to examine this effect by comparing valuations of generic health states with valuations of disease specific health states by a laymen panel.

Methods

Participants

From a large population sample of 560 lay people, a sample of 126 people were selected, that was representative for the Dutch population regarding age, sex, education and disease experience.

Health state description

For the description of the generic health state, we used and extended version of the EQ-5D classification system. This extended version of the EQ-5D classification system, the EQ-5D+, describes health with five levels of severity in six dimensions: mobility, self-care, usual activities, pain/discomfort, anxiety/depression, and cognition [8]. We presented 19 EQ-5D health profiles for valuation. These health profiles ranged from mild to severe. The disease specific health states consisted of an EQ-5D+ health profile and additional disease specific information, like a formalized clinical description of the disease and a visual aid to indicate the location of the condition on the body. In cases of altered physical appearance of the patient due to the condition, we added a picture of the condition to the vignette. In total, we presented 23 disease specific health state descriptions for valuation.

Valuation technique

To assess the health state valuations of the generic and disease specific health states, we used the Visual Analogue Scale (VAS) method conform the EQ-5D format. The VAS requires participants to score the health state on a vertical thermometer graded from 0 (worst imaginable health state) to 100 (best imaginable health state).

Valuation procedure

To assess the health state valuations of the generic and disease specific health states, we
This study was part of the Mild Diseases and Ailments Study [9, 10]. The Mild Diseases and Ailments Study followed a two step judgement procedure, consisting of a panel meeting and a questionnaire. For each of the panel meetings, 18 participants were invited. During the three-hour panel meeting, the purpose of the study, the EQ-5D classification system and the VAS method were explained. Subsequently, the participants valued ten disease specific health states with the VAS, including four disease specific health states from this study. During the valuation of the health states, the participants wrote down their value on a white board and explained how they reached their valuation. The participants were allowed to change their valuations at any time. Three weeks after the panel meeting, the participants received a questionnaire by mail. In this questionnaire, the participants were asked to value 22 disease specific health states and 6 generic health

states. We developed ten versions of the questionnaire, which were randomly assigned to the participants.

Analysis

The following formula was used to determine the VAS values for the health states:

VAS value = 1 – (VAS/100).

We calculated mean VAS values and standard deviation for each disease specific and generic health state. The difference between mean VAS values of generic and disease specific health state with the same EQ-5D+ profile were also calculated to establish the effect of the disease label. In order to examine the concordance between mean VAS values of the disease specific and generic health states, we calculated the Spearman correlation coefficient. We did not exclude or transform any data used for the calculation of the VAS values.

Results

Of the 126 lay people that were selected, a total of 105 agreed to participate in this study. Half of the respondents was male and the average age of the respondents was 49 years old. 99% (n = 104) of the participants that attended the panel meeting completed the questionnaire. It took the participants on average 1 hour and 56 minutes to complete the questionnaire.

Table 4.1 presents the mean VAS values the participants assigned to the generic and disease specific health states. Regarding the severity of the condition, indicated by the sum score of the EQ-5D+ profile, the participants ranked the generic and disease specific health states logically. High VAS values were assigned to severe health states and low VAS values to mild health states. The standard deviations of the VAS values of generic and disease specific health states did not differ, indicating a similar precision. In addition, the Spearman correlation coefficient between mean generic and disease specific VAS values is 0.94, indicating a similar ranking of the health states.

Regarding 19 of the 23 health states, mean disease specific VAS values were numerically higher ('more severe') compared to mean generic VAS values ($p < 0.001$). The difference in VAS values between generic and disease specific health states ranged from -0.08 to 0.28. The mean difference of the VAS values was 0.10. Especially conditions characterized by altered physical appearance had numerically higher VAS values (range 0.20 to 0.72) compared to similar generic health states (range 0.05 to 0.71). Furthermore, difference in mean VAS values was greater for less severe diseases. For mild health states, the mean difference in VAS values was 0.17, whereas for severe health states the mean difference in VAS values was 0.02.

Table 4.1 Difference in mean VAS values of disease specific and generic health states (HS)

		Disease specific HS		Generic HS		
Disease label	EQ-5D+	n	mean	n	mean	Δ
Allergic rhinitis	111211	34	0.27	14	0.05	0.22
Eczema	111211	34	0.27	14	0.05	0.22
Onychomycosis	111211	103	0.06	14	0.05	0.01
Tinea pedis	111211	34	0.14	14	0.05	0.09
Acne	111211	34	0.20	17	0.08	0.12
Eczema	111211	103	0.36	17	0.08	0.28
Stomach complaints	112311	34	0.42	17	0.19	0.23
Menopausal symptoms	112321	34	0.33	18	0.20	0.14
Eczema	112412	34	0.47	16	0.51	-0.03
Vertebral fracture	333311	34	0.66	17	0.43	0.23
Multitrauma, acute phase	333221	34	0.55	18	0.45	0.10
Larg burn, stable phase	333321	34	0.69	14	0.47	0.22
CVA	333323	103	0.80	17	0.73	0.06
Severe brain injury	224333	34	0.80	16	0.60	0.21
Large burn, acute phase	334431	34	0.72	14	0.71	0.01
Paraplegia, stable phase	544321	103	0.79	17	0.77	0.03
Multitrauma, stable	334333	34	0.76	14	0.70	0.06
Multitrauma, acute phase	445431	34	0.76	16	0.78	-0.02
Brain injury, severe, acute	335334	34	0.90	16	0.77	0.13
Paraplegia, acute phase	544431	34	0.81	18	0.89	-0.08
Quadriplegia, stable phase	555231	34	0.90	14	0.83	0.07
Quadriplegia, acute phase	555431	34	0.87	16	0.86	0.01
Multitrauma, acute phase	445434	34	0.86	16	0.91	-0.05

* 0 =full health, 1 = worst imaginable health state

Discussion

In this study, we found that the panel of lay people assigned numerically higher ('more severe') VAS values to disease specific health states compared with generic health states. The precision of the VAS values assigned to generic and disease specific health states was similar. This suggests that adding disease specific information to the health state description causes a shift in the VAS values.

It is inevitable that describing health with a generic description system results in information loss. With the six dimensions of the EQ-5D+ classification system, disease specific symptoms like altered physical appearance cannot be described. However, altered physical appearance is a prominent symptom of certain diseases and injuries and it

imposes a significant burden upon the patient [11, 12]. Several studies showed that patients with an altered physical appearance have a significantly decreased health-related quality of life [12-14]. For health state valuations information on altered physical appearance of the patient proves to be equally important. The results of this study showed that for conditions characterized by altered physical appearance, mean difference in VAS values between generic and disease specific health states increased almost twofold compared with mean difference in VAS values of the other health states.
Surprisingly, we found a greater difference in VAS values between minor generic and disease specific health states. This indicates that especially in minor diseases, the disease specific description contains information on prominent symptoms not reflected in the generic description. The minor diseases as well as their consequences may be easy to imagine for lay participants, especially since the minor disease were common, occurring frequently in the Dutch population. The severe diseases valued in this study occur less frequently and are therefore less familiar to lay people. This unfamiliarity in addition to the difficulty for healthy participants to imagine the impact of the severe conditions may have caused the participants to particularly use the added EQ-5D+ data in their valuation of the disease specific health states. In contrast, in the valuations of minor disease specific health states, the laymen panel may have valued the disease specific symptoms rather than the added EQ-5D+ data, causing a greater difference in VAS values between minor generic and disease specific health states.
The difference in VAS values assigned to minor generic and disease specific health states may also be explained by the valuation technique. In this study, we used the VAS to elicit the preferences of lay people for the generic and disease specific health states. However, the VAS is not choice-based and is therefore argued to be unable to capture the strength of the preferences for health states [15]. In contrast to the VAS, choice-based valuation techniques, like the time trade-off, have a threshold below which participants refuse to trade-off anything. The VAS lacks such a threshold and consequently the values that participants assign to minor health states with the VAS may overestimate the difference between minor conditions. We assume that the differences in values assigned to minor generic and disease specific health states are smaller when preferences are elicited with a choice-based valuation technique. This, however, remains to be investigated.

In summary, we conclude that lay people value disease specific health states as more severe compared to generic health states. The difference in health state valuation is greater with minor health states and health states that affect the physical appearance of the patient. This indicates that disease specific health state descriptions that contain information about prominent symptoms not reflected in the generic health state cause a shift in VAS health state valuations.

References

1. Field, M.J. and M.R. Gold, Summarising population health: Directions for the development and application of population health metrics. . 1998, Washington D.C.: Institute of Medicine: National Academy Press.
2. Murray, C.J. and A.K. Acharya, Understanding DALYs (disability-adjusted life years). J Health Econ, 1997. 16(6): p. 703-30.
3. Essink-Bot, M.L. and G.J. Bonsel, How to derive disability weights?, in Summary measures of population health: concepts, ethics, measurement and applications., C.J.L. Murray, A.D. Lopez, and J.A. Salomon, Editors. 2002, World Health Organization: Geneva.
4. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group Health Policy, 1990. 16(3): p. 199-208.
5. Rabin, R. and F. de Charro, EQ-5D: a measure of health status from the EuroQol Group. Ann Med, 2001. 33(5): p. 337-43.
6. Patrick, D.L. and R.A. Deyo, Generic and disease-specific measures in assessing health status and quality of life. Med Care, 1989. 27(3 Suppl): p. S217-32.
7. Stouthard, M.E., M.L. Essink-Bot, and G.J. Bonsel, Disability weights for diseases. A modified protocol and results for a Western European region. European Journal of Public Health, 2000. 10: p. 24-30.
8. Krabbe, P.F., et al., The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. J Clin Epidemiol, 1999. 52(4): p. 293-301.
9. Janssen, M.F., E. Birnie, and G. Bonsel, Feasibility and reliability of the annual profile method for deriving QALYs for short-term health conditions. Med Decis Making, 2008. 28(4): p. 500-10.
10. Bonsel, G.J., M.F. Janssen, and E. Birnie, Mild diseases & Ailments Study. Ziektelast-bepaling van 27 vermoedelijk 'lichte' ziekten ten behoeve van rationele beperking van het geneesmiddelenpakket. 2004, University of Amsterdam, Academic Medical Center: Amsterdam.
11. Newell, R.J., Altered body image: a fear-avoidance model of psycho-social difficulties following disfigurement. J Adv Nurs, 1999. 30(5): p. 1230-8.
12. Rumsey, N., et al., Altered body image: appearance-related concerns of people with visible disfigurement. J Adv Nurs, 2004. 48(5): p. 443-53.
13. Fauerbach, J.A., et al., Effect of early body image dissatisfaction on subsequent psychological and physical adjustment after disfiguring injury. Psychosom Med, 2000. 62(4): p. 576-82.
14. Lubeck, D.P., Measuring health-related quality of life in onychomycosis. J Am Acad Dermatol, 1998. 38(5 Pt 3): p. S64-8.
15. Brazier, J., M. Deverill, and C. Green, A review of the use of health status measures in economic evaluation. J Health Serv Res Policy, 1999. 4(3): p. 174-84.

5

The measurement of long-term health-related quality of life after injury: comparison of EQ-5D and the health utilities index

Suzanne Polinder, Juanita A. Haagsma, Gouke J. Bonsel, Marie-Louise Essink-Bot, Hidde Toet, Ed F. Van Beeck

Injury Prevention. 2010; 16: 147-153

Abstract

Objective Empirical head-to-head comparison of the health utility index (HUI) 2 and 3 and the EuroQol-5D (EQ-5D) in injury patients of all severity levels to obtain more insight in the strengths and limitations of the multi-attribute utility measures (MAUI) to estimate utility losses in injury populations.

Design and patients A self-assessment survey that included the EQ-5D, HUI2, and HUI3 to measure generic health-related quality of life.

Patients Injury patients in The Netherlands 2 years after they attended the emergency department.

Main outcome measures Shannon's index and Shannon's evenness index were used to assess absolute and relative informativity, both for the summary scores and by dimension. The study also analyzed convergent and construct validity of the MAUI.

Results Mean summary scores significantly differed between the instruments, with highest summary scores for HUI2 (0.88), followed by HUI3 (0.80) and EQ-5D (0.78). Absolute and relative informativity by dimension was highest for the HUI3 descriptive system. The HUI3 was most sensitive for ageing and comorbidity. The largest differences between the MAUI were found for pain/discomfort and anxiety/depression/emotion. The largest differences in discriminative power between the MAUI were seen for skull-brain injury, internal organ injury and upper extremity fractures.

Conclusions Different MAUI resulted in significantly different summary scores. The instruments and their dimensions performed differently for injury severity levels, ageing, comorbidity, and injury groups. A combination of the HUI and EQ-5D should be used in studies on injury-related disability, because the combination covers all relevant health dimensions, is applicable in all kinds of injury populations and in widely different age ranges.

Introduction

Health-related quality-of-life (HRQoL) has become an important consideration in the allocation of health care resources [1]. In the field of injury prevention, policy decisions could be supported by metrics on HRQol and also disability [2]. However, this is hampered by a lack of consensus on the preferred methods to arrive at these metrics, as can be derived from the wide variety of different approaches that are used by injury researchers [2, 3]. Therefore, in a first attempt to reduce the heterogeneity of applied methods, a European group published guidelines for the conduction of follow-up studies measuring injury-related disability [3]. Here it is advised to use a combination of the EuroQol-5D (EQ-5D) and the Health Utilities Index (HUI) 3 in all studies on injury-related disability as common core. The HUI (mark 2 and 3) and the EQ-5D are frequently used generic HRQoL measures [4-6] and aim to cover the full spectrum of disease and disability.

Both the HUI and the EQ-5D are multi-attribute utility instruments (MAUI), which are standardized health state classifications that can be used to obtain a single summary score (utility score) or so-called preference weight for different health states, based on preferences of the general public [1, 7]. At the core of any MAUI is a classification system consisting of multiple attributes (dimensions) with ordered levels for each dimension [8]. In this way, the health status of patients with problems on several dimensions can be quantified in a single metric between 0 (worst possible health status, and some health states are allocated index scores worse than dead) and 1 (best possible health status). This provides opportunities to compare the health status of patient groups with different diseases (e.g. heart disease versus cancer) and injury types (e.g. skull-brain injuries versus hip fractures).

The HUI and EQ-5D have been found to be acceptable, feasible, valid, and reliable in several population and patient studies [9-12]. Clear differences in the HUI and EQ-5D exist in definitions of health, inclusion of health dimensions, and construction of the formula. The literature has shown variation in summary scores between EQ-5D and the HUI for similar health states [9, 11, 13, 14]. These differences have the undesirable effect that the distinct instruments yield different utilities for similar health states. Furthermore, it is not yet clear which method provides the most valid summary scores within comprehensive samples of injury patients.

It seems generally believed that the HUI is a more responsive utility measure than the EQ-5D because of the crude level structure of the EQ-5D compared to the HUI. However, the performance characteristics of an instrument may be population specific [15], and the question arises as to whether this also holds for the heterogeneous group of injury patients. To be able to interpret the ability to discriminate between different injury and patient groups and different severity levels, a head-to-head comparison between the HUI2, HUI3 and EQ-5D is needed. As far as we know, published evidence of head-to-head comparisons between these three MAUIs among injury patients have not yet been conducted [3]. A comparison is needed to obtain more insight into the strengths and

limitations of both methods to estimate utility losses in injury populations. This could support further consensus development on preferred methodologies within the injury research field.

This paper describes a large follow-up study in which functional outcome was assessed with the EQ-5D and the HUI 2 years after injury in a comprehensive population of hospitalized and non-hospitalized injury patients. We will address the following questions:

1) Are there differences in summary scores between the EQ-5D and the HUI (mark 2 and 3) in an injury population sample?;
2) What is the discriminative power of the three systems in terms of their ability to distinguish between different levels of HRQoL among injury patients of all external causes and severity levels?;
3) What is the discriminatory power of the three systems based on head-to-head comparisons assessed with Shannon's indices of informativity?

Methods

Study population

We conducted a patient follow-up study among injury patients aged 15 years and older, who had visited one of the emergency departments of the Dutch Injury Surveillance System (LIS) [16, 17]. All unintentional and intentional injuries are recorded. LIS has been implemented in 17 hospitals in The Netherlands (15% coverage), which are considered to be representative for the total population. We used the data of the 1781 respondents on the 24-month questionnaire. The questionnaire was designed to collect information on functional outcome, sociodemographic and injury-related characteristics and healthcare use. A non-response analysis was performed by multivariate logistic regression. The study sample was stratified by type of injury (39 injury groups [18]) and admission so that severe, less common injuries were overrepresented. Data were corrected for non-response and sample stratification [16, 17]. Only patients with full information on all three HRQoL measures were included (n=1285).

International availability of comparable data

The questionnaire included two generic quality of life measures: the HUI mark 2/3 [4] and the EQ-5D. Table 5.1 gives an overview of the instrument properties.

Table 5.1 Overview of HRQoL instrument properties

	Dimensions/health domains/attributes	Nr. of possible health states	Valuation technique	Boundaries
HUI2	Sensation (vision, hearing, speech), mobility, emotion, cognition, self-care, pain	24,000	Standard gamble	-0.03 to 1.00
HUI3	Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain	972,000	Standard gamble	-0.36 to 1.00
EQ-5D	Mobility, usual activities, self-care, pain, anxiety/depression	243	Time trade-off	-0.59 to 1.00

The EQ-5D self-classifier comprises fives dimensions of HRQoL (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) each with three levels (for instance: no problems, some/moderate problems and extreme/unable to) [6]. Accordingly, the EQ-5D classification system distinguishes 243 different health states. The EQ-5D was analyzed using the so-called York A1 tariff, which is based on time trade-off preferences from the general public of the UK [5]. Reference scores for the EQ-5D index were obtained from the York study [19]. The EQ-5D is well able to describe a heterogeneous injury population and to discriminate among specific injuries [16]. Moreover, the EQ-5D has been recommended for (economic) evaluation of trauma care at a consensus conference [20]. Because the EQ-5D classification does not include memory patterns and/or ability to concentrate, an item was added on cognitive ability [21]. The EQ-5D supplemented by the cognitive dimension is referred to as the EQ-6D or the EQ-5D+.

The HUI is a self-administered health-status questionnaire consisting of 15 questions, classifying respondents into either the HUI2 or HUI3 health states. Responses to the questionnaire are converted by an algorithm [22] to levels in the complementary HUI2 and HUI3 health status classification systems [4] to form seven and eight-element health-state vectors, respectively. From these vectors, single-attribute and overall health state summary scores are calculated using the respective HUI2 and HUI3 utility functions [4, 23, 24], which are based on the standard gamble method, using preferences from the general Canadian population [23].

Analysis

Descriptive statistics were used to compare the mean summary scores, standard deviations and the number of missing cases per item. Missing values were defined as those cases where no answer was provided. Floor or ceiling effects are considered to be present if more than 15% of respondents achieved the lowest or highest possible score,

respectively [25]. Questionnaires should exhibit minimal floor and ceiling effects to be optimally able to detect difference and change.

To assess convergent validity, the extent to which the three instruments measure the same concepts, paired comparisons of the mean scores across measures were tested with the Paired Student's t test.

In the absence of a gold standard to measure health state utility, there is no clear technique to determine the construct validity of utility measures. A way to examine the construct validity is to examine whether summary scores are different for distinctive groups following a priori hypotheses of the expected patterns (sizes and directions) of the differences (known groups validity) [26]. Comparisons were made between EQ-5D, HUI2 and HUI3 by age, sex, injury group, and different severity levels (multiple injury, comorbidity). Differences regarding the mean summary scores for EQ-5D, HUI2 and HUI3 between groups were tested with a one-way analysis of variance.

The Shannon index and the Shannon evenness index of informativity were used to assess discriminatory power of each classification system [27]. The methodology of Shannon indices originates from the field of information theory, but can in principle be applied to any classification, including health classification systems or MAUI such as the EQ-5D, HUI, and the short form health survey (SF-6D) [8]. Shannon indices can be calculated by dimension separately or by MAUI as a whole. The basic characteristic of Shannon's indices is explained as follows. In an item with two response categories in which one response category has a very high (or low) endorsement, for example, more than 0.95 (or less than 0.05), response category transmits very little information because one can predict with more than 95% certainty in what response category the answer will be [28]. Conversely, in case of an even distribution, the health dimension is being most efficiently used, which means that the discriminant ability of the level descriptors is maximal. This characteristic of an even distribution underlies Shannon's indices. The Shannon index (H') combines the number of non-empty categories defined by a system, and measures to what extent the information is (empirically) evenly spread over the non-empty categories. The higher the index H' is, the more information is captured by the system. The Shannon Evenness index (J') exclusively reflects the evenness (rectangularity) of a distribution, regardless of the number of levels. Five dimensions allowed head-to-head comparison of informativity: mobility/ambulation; anxiety/depression/emotion; pain/discomfort (EQ-5D; HUI2; HUI3); self-care (EQ-5D; HUI2); and cognition (EQ-6D; HUI2; HUI3). To calculate Shannon's indices by instrument as a whole and by injury group, permutations are treated as unique categories (e.g. 243 categories for EQ-5D). As the number of observations in our study (n=1285) is lower than the number of theoretically possible permutations in HUI2 (24000) and HUI3 (972000), maximum informativity cannot be reached a priori. The basic characteristics of the Shannon's index are described elsewhere [8]. Regarding the Shannon indices, the observed number of health states in a population are used, not the theoretical possible number of health states.

Results

Descriptive comparison of questionnaires

Of the 1781 injury patients who responded to the 2-year questionnaire 1561, 1541, and 1454 persons fully completed the EQ-5D, HUI2, and HUI3, respectively (table 5.2). Mean summary scores for the injury population were highest for the HUI2. Generally, the questionnaires demonstrated no floor effects. All three questionnaires showed ceiling effects; 40% of the population indicated to have no problems on EQ-5D, against 25% on HUI2 and 24% on HUI3. The HUI2 and HUI3 have similar interquartile ranges: 0.15 and 0.18, respectively; conversely, the interquartile range of the EQ-5D was wider (0.23).

Table 5.2 Descriptive statistics of the EQ-5D and HUI

	HUI2	HUI3	EQ-5D
Unweighted data (n=1,285[1])			
Mean	0.88	0.80	0.78
Standard deviation	0.13	0.22	0.20
Median	0.92	0.87	0.80
Interquartile range	0.12	0.28	0.24
% patients minimum possible score	0	0	0
% patients maximum possible score	18.3	14.9	22.2
Weighted data[2]			
Mean	0.91	0.85	0.86
Standard deviation	0.11	0.19	0.16
Median	0.95	0.92	0.92
Interquartile range	0.15	0.18	0.23
% patients minimum possible score	0	0	0
% patients maximum possible score	25.2	23.5	40.2

[1] Only patients that responded on all three HRQOL measures were included. In total 1,561, 1,541, 1,454 patients responded on the EQ-5D, HUI2, and HUI3 questionnaire, respectively.

[2] Data are corrected for non-response and stratification and are representative for a population of injured patients who visited an ED in the Netherlands.

Comparison of summary scores

Following the results from the paired t-test on the unweighted data, the mean utilities derived with the three MAUI were significantly different from each other (EQ-5D vs HUI2 t=-82.9; EQ-5D vs HUI3 t=18.7; HUI3 vs HUI2 t=-122.0). In table 5.3 the differences in summary scores for specific determinants were analyzed by MAUI.

Table 5.3 HUI2, HUI3, EQ5D summary scores by determinants (15 years and older, hospitalized) and the relationship between determinants and the MAUIs

	HUI2	HUI3	EQ-5D
Total	0.86	0.77	0.75
Age **			
15-24	0.92	0.86	0.80
25-44	0.89	0.83	0.78
45-64	0.85	0.75	0.70
65-74	0.84	0.70	0.74
75-84	0.79	0.60	0.71
85+	0.63	0.42	0.64
Comorbidity **			
No	0.89	0.82	0.77
1	0.81	0.69	0.71
>=2	0.76	0.58	0.67
Hospitalization **			
1-3 days	0.89	0.78	0.74
4+ days	0.84	0.72	0.67
Type of injury **			
Type of injury **			
Skull-brain injury	0.84	0.72	0.80
Facial fracture, eye injury	0.88	0.79	0.80
Spine, vertebrae	0.81	0.68	0.63
Internal organ injury	0.88	0.81	0.81
Upper extremity fracture	0.90	0.82	0.77
Upper extremity, other	0.93	0.84	0.83
Hip fracture	0.79	0.62	0.66
Lower extremity fracture	0.87	0.78	0.67
Lower extremity, other	0.88	0.82	0.71
Superficial injury, open wounds	0.90	0.81	0.79
Other injury[1]	0.88	0.77	0.84

Data are corrected for non-response and stratification and are representative for a population of injured patients who visited an ED in the Netherlands. Comparison (using ANOVA) of mean values: ** =p-value < 0.001.

[1] Other injury: consists of whiplash, poisoning, burns, injury of nerves.

For several determinants, the differences in summary scores between the instruments were not in the same direction as the total summary scores (HUI2>HUI3>EQ-5D). The HUI3 scores were relatively low for elderly patients (age over 65 years), with a more than 0.20 utility difference compared to the HUI2 and EQ-5D for the age group over 85 years. Furthermore, the HUI3 was most sensitive for the presence of comorbidity. Patients with two or more comorbid conditions had a lower summary score (0.19 lower compared to the mean score) for the HUI3, compared with a reduction of 0.10 and 0.08 for the HUI2 and EQ-5D, respectively.
All three instruments showed significant differences in summary score by length of stay in hospital. The HUI2 scores were higher than HUI3 and EQ-5D for all types of injury. The HUI3 showed lower summary scores than the EQ-5D for skull-brain injury, facial fractures and hip fracture.

Informativity

Figure 5.1 shows absolute informativity (Shannon's H') and relative informativity (Shannon's evenness J') of the common dimensions among the three instruments. Absolute informativity (H'), or diversity (the degree to which health states were distributed equally among the injury patients), was highest for HUI3 in most dimensions, with largest differences between HUI3 and EQ-5D in the dimensions pain/discomfort (1.70 vs 1.02) and anxiety/depression/emotion (1.34 vs 0.64). Furthermore, for cognition large differences in absolute informativity were found between the HUI3 and EQ-6D (1.17 vs 0.63). HUI3 showed the highest relative informativity (J'), or evenness (the degree to which the instrument reflected the maximal diversity that was possible given the number of health states observed) for pain, emotion and cognition, with the largest differences with the other two instruments in the dimensions anxiety/depression/emotion (0.12 difference compared to EQ-5D; 0.05 difference compared to HUI2) and cognition (0.08 difference compared to HUI2; 0.05 difference compared to EQ-6D). For the dimension mobility/ambulation both the absolute and relative informativity were highest for EQ-5D in comparison with HUI2 and HUI3.
Table 5.4 shows Shannon's indices by classification system as a whole. Absolute informativity (H') was highest for HUI3 (6.08), followed by HUI2 (4.75), and lowest for EQ-5D (2.71). This means that most information was captured by the HUI3 classification system. Relative informativity (J') was highest for HUI3 (0.72) (0.19 higher than EQ-5D; 0.06 higher than HUI2). The EQ-5D, HUI2 and HUI3 descriptive systems distinguished 36, 150. and 347 observed different unique health states, accounting for 14.8%, 0.6% and 0.04% of all possible permutations, respectively.
The biggest differences in absolute informativity (Shannon's H') between EQ-5D compared to HUI2 and HUI3 was seen for skull-brain injury (2.83 vs 4.58 and 5.46, respectively) and internal organ injury (2.31 vs 4.37 and 4.65, respectively). Furthermore, a large difference in discriminative power exists between the EQ-5D and HUI3 for upper extremity fractures (2.33 vs 5.14).

For all MAUIs the discriminative power was higher for persons with comorbidity. Both the absolute and relative informativity were higher for the HUI2 and HUI3 compared to the EQ-5D. Also, for all MAUIs the absolute informativity (Shannon's H') was highest for long-term admitted patients. This implies that the observed health states among injury patients that were long-term admitted and/or with a comorbid disease provide the best reflection of the maximal diversity given the possible health states.

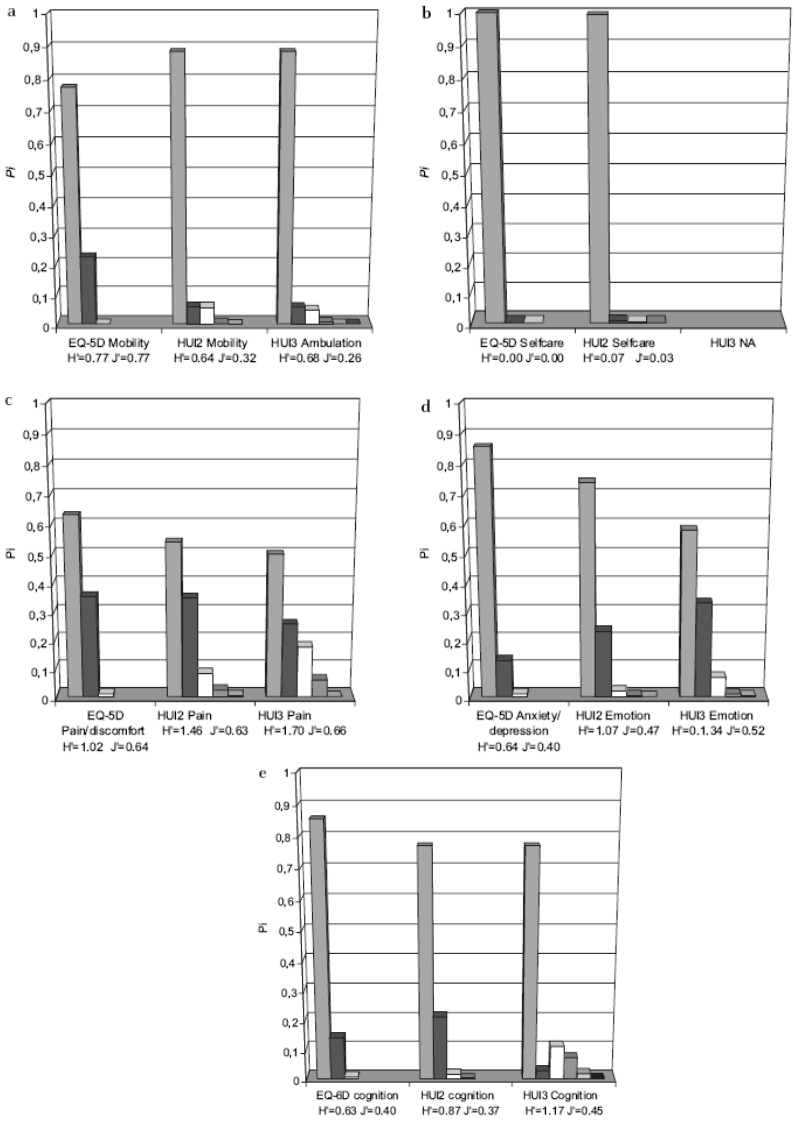

Figure 5.1 The Shannon Index and the Shannon Evenness Index for the common dimensions between EQ-5D, HUI2 and HUI3: comparison by dimension

Table 5.4 Shannon's index (*H'*) and Shannon's evenness index (*J'*) for EQ-5D, HUI2 and HUI3 by type of injury: comparison by instrument

	HUI2		HUI3		EQ-5D	
Nr of possible permutations	24,000		972,000		243	
Observed health states	150		347		36	
H' max[1]	14.55		19.89		7.93	
	H'	*J'*	*H'*	*J'*	*H'*	*J'*
Total injury population	4.75	0.66	6.08	0.72	2.71	0.53
Type of injury						
Skull-brain injury	4.47	0.62	5.34	0.63	2.55	0.49
Facial fracture, eye injury	3.42	0.47	3.86	0.46	1.71	0.33
Spine, vertebrae	4.12	0.57	4.77	0.57	2.91	0.56
Internal organ injury	4.37	0.60	4.65	0.55	2.31	0.45
Upper extremity fracture	3.98	0.55	5.14	0.61	2.33	0.45
Upper extremity, other	3.57	0.49	4.59	0.54	2.39	0.46
Hip fracture	4.58	0.63	5.46	0.65	2.83	0.55
Lower extremity fracture	4.36	0.60	5.19	0.61	2.81	0.54
Lower extremity, other	4.03	0.56	4.58	0.54	2.46	0.48
Sup. injury, open wounds	3.97	0.55	4.87	0.58	2.27	0.44
Other injury[2]	3.70	0.51	4.09	0.48	2.59	0.50
Comorbidity						
No	4.09	0.57	5.22	0.62	2.29	0.44
1	5.32	0.74	6.52	0.77	3.24	0.63
>=2	5.25	0.36	5.97	0.71	3.35	0.42
Hospitalization						
Not admitted	4.29	0.59	5.25	0.62	2.21	0.43
1-3 days	4.38	0.61	5.53	0.66	2.60	0.50
4+ days	5.11	0.71	6.26	0.74	3.28	0.63

[1] *H'* max: in case of a homogeneous distribution, the optimal amount of information is captured and *H'* has reached its maximum (*H'* max). Since the number of observations in our study (n=1,285) is lower than the number of theoretically possible permutations in HUI2 (24,000) and HUI3 (972,000), maximum informativity cannot be reached a priori.

[2] Other injury: consists of whiplash, poisoning, burns, injury of nerves.

Discussion

This paper focused on the ability of the HUI2, HUI3 and EQ-5D to discriminate between different levels of HRQoL among injury patients of all severity levels. Mean summary scores for the injury population were significantly different between the instruments, with highest summary scores for the HUI2 and lowest for EQ-5D. The HUI3 is most sensitive for HRQoL resulting from old age (over 65 years) and comorbidity. All three instruments demonstrated sensitivity for differences in injury type and hospitalization; in addition, they showed similar rankings between injury patient groups. Absolute and relative informativity by dimension was highest for the HUI3 descriptive system. The largest differences between the MAUI were found for pain/discomfort and anxiety/depression/emotion. EQ-5D appears to underperform in these two dimensions. In addition, EQ-6D appeared to have low discriminative power compared with HUI also on the added cognitive dimension. The biggest differences between the MAU in discriminative power between EQ-5D compared with HUI2 and HUI3 were seen for skull-brain injury, internal organ injury (HUI2) and upper extremity fractures (HUI3).

To the best of our knowledge, only a few studies have been published comparing the HUI and EQ-5D extensively [8, 9, 14, 29, 30]. Generally, the results of most of these studies agreed with ours in that there were significantly different utility values generated for the MAUIs. As far as we know, published evidence of head-to-head comparisons between the HUI and the EQ-5D among injury patients have not yet been conducted [3].

Typical for injuries are their heterogeneous functional sequelae and recovery patterns. Therefore it is important that the MAUI used has good discriminative power for the severity of injury. All three instruments showed significant differences of summary scores by injury group, with relatively low scores for injuries of the spine/vertebrae and hip fractures and high scores for superficial injury and upper extremity injury, close to the health state of the Dutch general population [31]).

The large differences found between the HUI2, HUI3 and EQ-5D may be confounded by the different valuation and/or scoring methods of the MAUI. First of all, there are important differences in the applied health state valuation technique. The HUI systems used the standard gamble as a valuation technique [32] and the EQ-5D used the time trade-off technique [9]. The time trade-off technique has been shown to yield lower results compared with the standard gamble technique [9, 33], which indicate that lower EQ-5D scores could be expected. Second, the scoring functions for the EQ-5D utility values were derived from samples of the UK population, which may differ from preferences given by those in Canada [5].

Furthermore, an advantage of the HUI3 is its potentially greater discriminatory power in a wide range of diseases because it defines 972000 unique health states, whereas the EQ-5D defines 243 [9]. We found that the EQ-5D showed highest ceiling effects: 40% of the population indicated to have no problems on EQ-5D, against 25% on HUI2 and 24% on HUI3. These high 'ceiling effects' are not surprising, because in earlier research it was

shown that a large part of the patients recover within 2 years from an injury [16]. However, in 16% of the cases the EQ-5D finds no disability where the HUI result in functional problems and 32% of indicated that they were not fully recovered two years after the injury.

Performance in terms of informativity of EQ-5D, HUI2 and HUI3 of the common dimensions varies over dimensions. Absolute and relative informativity by dimension was highest for the HUI3 descriptive system. Shannon's indexes 'translated' this difference adequately. Apparently, the EQ-5D would benefit from more levels on the pain/discomfort dimension. Regarding the cognition dimension, the difference in absolute informativity between HUI2, EQ-6D and HUI3 might be explained by the extra levels in HUI3, but the higher J' value in HUI3 suggests an alternative contributive factor. The very low informativity in self-care among all instruments might be explained by adaptation of skills to the new situation, even when persons still experience health problems two years after the injury.

All three measures have the advantage that they are available in formats designed for self-completion, and include a comprehensive health status classification system. The HUI with 15 questions is somewhat more elaborate than the EQ-5D, which consists of five questions. However, fewer questions may yield a higher response rate and fewer missing scores in a mail survey. In our research we also found that more persons fully completed the EQ-5D compared to the HUI2 and HUI3, but the differences were modest (88% versus 87% and 82%, respectively).

Most information is captured by the HUI3 classification system. However, it should be considered that this also results in the fact that the HUI3 is very sensitive for HRQoL reduction due to comorbidity and ageing. For patients in our study above the age of 75 years and with one or more comorbidities (n=50), the summary score of the HUI3 is 0.55, which is much lower than for the HUI2 and EQ-5D (0.75 and 0.72, respectively). This could mainly reflect the influence of other health problems instead of injury consequences, and researchers should be cautious with the HUI3 as a stand-alone measure in injured elderly populations.

This study showed that different MAUI resulted in significantly different summary scores. Furthermore, the HUI2, HUI3 and EQ-5D and their dimensions perform differently for injury type, hospitalization and length of stay in hospital, comorbidity, and ageing. These differences have the undesirable effect that the distinct instruments yield different utilities for similar health states. To demonstrate the functional outcome of injury patients, clinicians and researchers should be aware of these differences between the MAUI.

Decisions about which HRQoL measure to use will be influenced by a range of factors. We showed that the HUI classification system is more informative than the EQ-5D, in particular for patients with skull-brain injury, internal organ injury and upper extremity fractures. Most information is captured by the HUI3 classification system, but this does not seem to have enormous advantages in practice. Nevertheless, the EQ-5D seems to be the utility measure more often applied in injury research in all injury populations and it

was recommended for economic evaluations. Although noting injury outcome limitations of the EQ-5D, such as the absence of a cognitive dimension, the EQ-5D being freely available, its simplicity and high response rates, and with many language versions available, seems suitable for inclusion in studies in the injury population.
Well-founded choices are essential with regard to the type of measure to be used for analyzing HRQoL injury patients. We advise to use a combination of the HUI and EQ-5D in studies on injury-related disability, in line with earlier published guidelines [3]. This combination covers all relevant health dimensions, is applicable in all kinds of injury populations and in widely different age ranges.

References

1. Gold, M.R., et al., Cost-effectiveness in health and medicine. 1996, New York: Oxford University Press.
2. Segui-Gomez, M. and E.J. Mackenzie, Measuring the Public Health impact of injuries. Epidemiol Rev, 2003. 25: p. 3-19.
3. Van Beeck, E.F., et al., Guidelines for the conduction of follow-up studies measuring injury-related disability. J Trauma, 2007. 62(2): p. 534-50.
4. Feeny, D., et al., Multi-attribute health status classification systems. Health Utilities Index. Pharmacoeconomics, 1995. 7(6): p. 490-502.
5. Dolan, P., Modeling valuations for EuroQol health states. Med Care, 1997. 35(11): p. 1095-108.
6. Brooks, R., EuroQol: the current state of play. Health Policy, 1996. 37: p. 53-72.
7. Drummond, M., et al., Methods for the Economic Evaluation of Health Care Programmes. 1997, Oxford, U.K.: Oxford Medical Publications.
8. Janssen, M.F., E. Birnie, and G.J. Bonsel, Evaluating the discriminatory power of EQ-5D, HUI2 and HUI3 in a US general population survey using Shannon's indices. Qual Life Res, 2007.
9. Bosch, J.L. and M.G. Hunink, Comparison of the Health Utilities Index Mark 3 (HUI3) and the EuroQol EQ-5D in patients treated for intermittent claudication. Qual Life Res, 2000. 9(6): p. 591-601.
10. Glick, H.A., et al., A comparison of preference assessment instruments used in a clinical trial: responses to the visual analog scale from the EuroQol EQ-5D and the Health Utilities Index. Med Decis Making, 1999. 19(3): p. 265-75.
11. Marra, C.A., et al., A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med, 2005. 60: p. 1571-1582.
12. Kopec, J.A. and K.D. Willison, A comparative review of four preference-weighted measures of health-related quality of life. J Clin Epidemiol, 2003. 56(4): p. 317-25.
13. Luo, N., et al., Self-reported health status of the general adult U.S. population as assessed by the EQ-5D and Health Utilities Index. Med Care, 2005. 43(11): p. 1078-86.
14. Oostenbrink, R., A.M. HA, and M.L. Essink-Bot, The EQ-5D and the Health Utilities Index for permanent sequelae after meningitis: a head-to-head comparison. J Clin Epidemiol, 2002. 55(8): p. 791-9.
15. Essink-Bot, M.L., et al., An empirical comparison of four generic health status measures. The Nottingham Health Profile, the Medical Outcomes Study 36-item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. Med Care, 1997. 35(5): p. 522-37.
16. Polinder, S., et al., Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma, 2007. 2007(62): p. 133-141.

17. Polinder, S., et al., Prevalence and prognostic factors of disability after childhood injury. Pediatrics, 2005. 116: p. 810-817.
18. Lyons, R.A., et al., Methodological issues in comparing injury incidence across countries. Int J Inj Contr Saf Promot, 2006. 13(2): p. 63-70.
19. Kind, P., G. Hardman, and S. Macran, UK population norms for EQ-5D. 1999, Centre of Health Economics, University of York: York.
20. Bouillon, B., et al., Quality of life in patients with multiple injuries--basic issues, assessment, and recommendations. Restor Neurol Neurosci, 2002. 20(3-4): p. 125-34.
21. Krabbe, P.F., et al., The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. J Clin Epidemiol, 1999. 52(4): p. 293-301.
22. Furlong, W., D. Feeny, and G.W. Torrance, Health Utilities Index (HUI©) Procedures Manual: algorithm for determining HUI Mark 2 (HUI2)/Mark 3 (HUI3) health status classification levels, health states, single-attribute utility scores and overall health-related quality of life utility scores, from 15-item self-complete health states questionnaires. 2000, Health Utilities Inc: Hamilton, Ontario.
23. Torrance, G.W., et al., Multi-attribute utility function for a comprehensive health status classification system: Health Utilities Index mark 2. Med Care, 1996. 34: p. 702-722.
24. Furlong, W., et al., Multiplicative, multi-attribute utility function for the Health Utilities Index Mark 3 (HUI3) system: a technical report. 1998, McMaster University Centre for healht Economics and Policy Analysis Working Paper 98-11: Hamilton, Canada.
25. Terwee, C.B., et al., Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol, 2007. 60(1): p. 34-42.
26. Streiner, D.L. and G.R. Norman, Health measurement scales. A practical guide to their development and use. 2003, New York; Oxford University Press.
27. Shannon, C.E., A mathematical theory of communication. The Bell System Technical Journal, 1948. 27: p. 379-423.
28. Janssen, M.F., et al., Comparing the standard EQ-5D three-level system with a five-level version. Value in Health, 2008. 11: p. 275-284.
29. Glick, H.A., et al., A comparison of preference assessment instruments used in a clinical trial: Resonpses to the visual analog scale from the EroQol EQ-5D annd the Health Utilities Index. Med Decis Making, 1999. 19: p. 265-275.
30. Marra, C.A., et al., A comparison of four indirect methods of assessing utility values in rheumatoid arthritis. Medical Care, 2004. 42(11): p. 1125-31.
31. Hoeymans, N., H. van Lindert, and G.P. Westert, The health status of the Dutch population as assessed by the EQ-6D. Qual Life Res, 2005. 14(3): p. 655-63.
32. Kopec, J.A. and K.D. Willison, A comparative review of four preference-weighted measures of health-related quality of life. J Clin Epidemiol, 2003. 56: p. 317-325.
33. Drummond, M.F., et al., Methods for the economic evaluation of health care programmes, 2nd edition. 1997, Oxford: Oxford University Press.

Part II

Comorbidity adjustment

6

Independent comorbidity and health-related quality of life of injury patients in the first year following injury: comparison of comorbidity adjustment approaches

Juanita A. Haagsma, Ed F. van Beeck, Suzanne Polinder, Hidde Toet, Gouke J. Bonsel

Submitted

Abstract

Introduction At present three approaches exist to deal with the impact of comorbidity in burden of disease studies; the maximum limit, additive and multiplicative approach. The aim of this study was to compare the three adjustment approaches in patients with an injury condition as primary condition and six comorbid conditions with non-trivial health impact.

Methods Health related quality of life data was assessed using EQ-5D data from a postal survey among 2,123 injury patients at 2½ and 9 months after attending the Emergency Department. Using data of injury patients with and without comorbidity who were restored from their injuries at 9-months follow-up, we compared the observed and predicted disability in comorbid cases. The predicted disability was calculated with the maximum limit, additive and multiplicative approach.

Results The disability of injury patients, measured with EQ-5D, increases significantly with the number of comorbid diseases. The three methods to account for comorbidity effects provided reasonable results. The adjustment approaches performed better in case of comorbid injury and severe chronic disease.

Conclusions Comorbidity has a high impact on disability measured with EQ-5D. Ignoring the effect of comorbid disease in burden of disease estimates restricts the use of disability adjusted life years in multi-morbid populations. Gains from care may easily be overestimated if a substantial number of patients suffer from additional conditions, which may lead to wrongful policy recommendations.

Introduction

Burden of disease studies quantify the health status of a population in order to facilitate policy makers in priority setting in health care and prevention [1, 2]. Commonly, their outcome is expressed in Disability Adjusted Life Years (DALY), a summary measure of population health. Apart from the obvious advantages of a uniform summary measure of population health, the calculation and interpretation of the burden of disease in terms of DALYs can be complicated if multiple conditions co-exist in individuals. The fact that multiple conditions co-exist in individuals may be a matter of chance, related to general susceptibility (e.g. advanced age) or the consequence of a disease with multiple systemic manifestations and remote complications. Nevertheless, assignment of the observed burden to separate conditions, either in descriptive terms or in terms of the total computed burden of disease, is arbitrary and several difficulties emerge.

A first difficulty is that the straightforward additive use of DALYs per disease is limited as this assumes that the total burden of two or more diseases is the sum of the burden of diseases taken separately. A second difficulty occurs in so-called counterfactual impact analysis of risk factors. What happens in terms of population DALYs if one disease is eradicated? The answer depends on a valid solution of the comorbidity assignment problem. Furthermore, the comparative outcomes research faces problems to the extent that outcome differences can virtually disappear through overriding effects of comorbid conditions. The common practice to exclude patients with comorbidity from participation in trials rather post-pones than solves the question concerning the average population effect of an intervention.

Comorbidity is defined as the presence of any clinical condition which qualifies for formal classification as a disease additional to the disease under study. Risk factors like advanced age, ethnic background or obesity are essentially not comorbid conditions, although principles described below might be applicable too. Comorbidity can be divided on biological grounds into at least three types: 1) unrelated (the most common case of two conditions happening by chance on the same individual), 2) indirectly related (through common risk factors while pathofysiology is grossly unrelated, in particular the pathway to symptoms), and 3) directly related (when pathofysiology shows that one condition can be regarded as natural consequences or parallel manifestation of the other condition) [3].

Regarding burden of disease studies at present three approaches exist to deal with the impact of comorbidity [4, 5]. These approaches are elaborated here, considering the previous three different types of commonality.

The first approach is the maximum limit (maxlimit) approach. This approach counts the disease with the highest overall disability weight. The approach assumes that a comorbid disease does not affect disability of a patient with a primary disease, unless the comorbid disease - in general terms - exceeds the disability of the former.

The second approach is the additive approach. This approach assumes that the additional effect (or more precise: utility loss) of comorbid disease simply adds to the effect (utility

loss) of the primary disease observed in uni-conditional patients. The disability weights of the comorbid diseases are added up.
The third approach is the multiplicative approach. This method assumes that a comorbid disease does increase the utility loss of a patient, though it is less than the sum of the utility loss of both diseases independently.
Here we present a systematic comparison of the three comorbidity adjustment approaches in patients with an injury as primary condition and common conditions with non-trivial health impact as the secondary condition.

Methods

Design

We compared the observed (gold standard) and predicted disability weight in comorbid cases, using EQ-5D data of injury patients with and without comorbidity. For this comparison, long-term follow-up data from injury patients could be used because the onset of injury is acute, causing immediate, yet usually temporary functional loss. This allows the measurement of utility loss due to comorbid injury and disease (comorbid utility loss) and utility loss due to the injury and disease separately (uni-conditional utility loss). The comorbid utility loss was used to calculate the observed disability weight in comorbid cases. The uni-conditional utility loss was used to calculate the predicted disability weight with the maximum limit, additive and multiplicative approach. See figure 6.1 for a schematic model of the design.

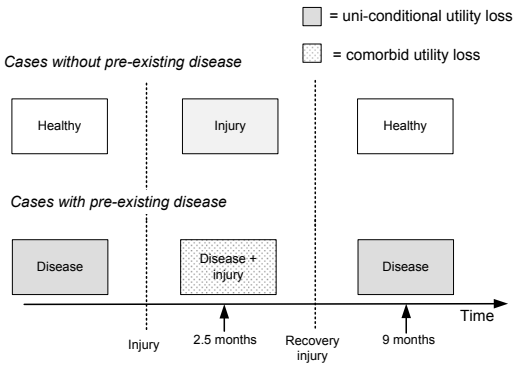

Figure 6.1 Design of the study

Patient data

The primary data sources were existing national registry data on injured patients at hospital admission, enhanced with functional outcome data obtained from patients by surveys at regular intervals. The registry data was derived from the Dutch Injury

Surveillance system, a permanent registry of injuries treated at the Emergency Department (ED) of a representative sample of 17 hospitals (about 10-15% of ED capacity) in the Netherlands. The registry collects information on age and sex of the patient, cause and type of the injury, body region affected and treatment of the sustained injury [6]. The follow-up survey was administered between 8 October 2001 and 31 December 2002 to a sample of 8,564 patients aged 15 years and older who attended the ED of the Dutch Injury Surveillance System [6]. The patients were treated at the ED followed by either hospital admission or direct discharge to the home environment. The sample of patients consisted of victims of traffic, home and leisure, occupational and sport accidents. The sustained injuries varied from minor to severe injury and hospitalized and non-hospitalized patients. The sample of patients was stratified, over sampling patients who were hospitalised. Each injury patient of the selected sample received a postal questionnaire 2½ months after the injury and 3,167 (37%) responded. The first questionnaire was made anonymous for privacy reasons. At 5, 9 and 24 months a follow-up questionnaire was sent to patients that responded to the preceding questionnaire. For these questionnaires the patients needed to give permissions by an informed consent form. The present study used a sample of 2,295 respondents (i.e. 27% of the original sample) that responded to the 2½ months and 9 month post-trauma survey.

Utility measurement and disability weight

To measure utility after injury, the questionnaire included the multi-utility attribute instrument EQ-5D. Regarding the EQ-5D classification system, subjects describe their health state by assigning themselves to one of three function levels (grades) in 5 separate domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression [7, 8]. Subsequently, the utility weight of that health state is computed by a formula that firstly yields a partial weight score for each domain depending on the reported level, and secondly adds the utility weights (also referred to as the tariff), which were derived at an earlier stage from preference data of the UK population [9]. We assumed burden of disease in terms of a disability weight (dw) to be represented by:

$$dw_t = 1 - u,$$

where u is the EQ-5D utility score at t months after sustaining the injury.

Observed disability weight in comorbid cases

The survey included a question that asked whether patients were restored from their injuries (yes/no). This study was restricted to injury patients who indicated that they were restored from their injury at 9 months follow-up. Furthermore, the questionnaire included 19 items regarding the presence of one or more chronic disease(s) prior to the injury to assess comorbidity [10]. Comorbidity is defined as the presence of any coexisting medical

conditions or disease processes additional to the injury that the injury patients sustained [11]. We selected six persisting diseases that were most often reported, i.e. chronic non-specific lung disease, heart disease, diabetes, backache, arthrosis and rheumatoid arthritis. The comorbid injury and disease disability weight was obtained from 2½ month EQ-5D data reported by injury patients with comorbid disease (c = 1) and it is represented by:

$dw_{injury + disease, t1} = (1 – u \mid c = 1)$,

where u is the EQ-5D utility score measured at 2½ month follow-up (t1) among injury patients with comorbid disease (c = 1).

Predicted disability weight in comorbid cases

To predict the disability weight in comorbid cases with the maximum limit, additive and multiplicative approach, uni-conditional disability weights were used.
The uni-conditional disability weight of the injury was obtained from the 2½ month EQ-5D data reported by injury patients without comorbid (c = 0) disease with the following formula:
$dw_{injury, t1} = (1 – u \mid c = 0)$,
where u is the EQ-5D utility score measured at 2½ month follow-up (t1) among injury patients without comorbid disease (c = 0).
The effects of the injury consequences on utility can be measured 2½ months after sustaining the injury, given that the selected patients were restored at 9 months post-injury. This assumption allows measurement of the disability effects of the comorbid disease only. Therefore, the uni-conditional disability weights of the disease were obtained from the 9-month EQ-5D data of injury patients with comorbid disease (c = 1) as follows:

$dw_{disease, t2} = (1 – u \mid c = 1)$,

where u is the EQ-5D utility score measured at 9 month follow-up (t2) among injury patients with comorbid disease (c = 1).

These uni-conditional disability weights were then used to calculate the predicted or disability weight in case of injury and comorbid disease according to the three approaches. In the examples supporting the formulas which were used to calculate the comorbid disability weights, the injury type leg fracture serves as the primary disease and heart disease as an example of the pre-existing disease.

Maximum limit approach
To calculate comorbid disability with the maximum limit approach, we used the following formula:

$dw_{combined} = dw_{leg\ fracture + heart\ disease} = \max(dw_{heart\ disease}, dw_{leg\ fracture})$.

The maximum limit approach is biased if two conditions affect different health domains of the EQ-5D (underestimation by model-based cancellation).

Additive approach
The additive approach comorbid disability weight was calculated as follows:

$dw_{leg\ fracture + heart\ disease} = dw_{leg\ fracture} + dw_{heart\ disease.}$

One limitation is that the combined disability may exceed 1.0. Bias may arise if both conditions affect the same health domain of the EQ-5D: the effect of diabetic foot amputation does not add to a lower leg fracture of the same leg (overestimation by data cancellation).

Multiplicative approach
The multiplicative approach comorbid disability weight was calculated with the formula:

$dw_{leg\ fracture + heart\ disease} = 1 - (1 - dw_{leg\ fracture}) * (1 - dw_{heart\ disease})$.

Analysis

For the analysis of the data the Statistical Package for the Social Sciences (SPSS) version 16.0 was used (SPSS inc, Chigaco, Ill). One-way ANOVA was used to test for differences in dis-ability weights between patients with and without comorbid disease. The Pearson correlation coefficient and intra class correlation (ICC) were used to test whether the distributions of the observed disability weights and the three model-predicted disability weights were correlated. To test whether the relations between observed and predicted disability weights depended on the severity of the comorbid disease, the pre-existing disease was grouped into two severity classes using the median value of the 9-month disability weight of patients with comorbid disease (0.21) as a cut-off (range 0.0-0.21 is less severe group).

Results

Of the 2,295 injury patients that completed the follow-up surveys 2½ and 9 months after they attended the Emergency Department, 1,036 (45.1%) indicated that they were restored from their injury at 9 months follow-up.

The 2½-month EQ-5D disability weights increased significantly with the number of comorbid conditions ($F=60.33$, $p<0.001$). For instance, injury patients with an upper extremity fracture and no comorbid disease had a mean EQ-5D disability weight of 0.10, whereas similar patients with two or more comorbid condition had mean EQ-5D disability weights of 0.20 and 0.23, respectively. Table 6.1 shows the uni-conditional disability weights of the injury groups that were obtained from 2½ month EQ-5D data of injury patients without comorbidity.

Table 6.1 Mean uni-conditional disability weights of the injury groups, obtained from EQ-5D data from injury patients at 2½ months follow-up

Injury type	N	Mean	CI[1]
Skull - brain injury	192	.10	.08−.12
Facial fracture, eye injury	51	.05	.03−.08
Spine, vertebrae	26	.24	.14−.34
Internal organ injury	61	.11	.07−.15
Upper extremity fracture	188	.10	.08−.13
Upper extremity, other injury	64	.11	.07−.15
Hip fracture	12	.29	.09−.49
Lower extremity fracture	151	.20	.17−.49
Lower extremity, other injury	74	.20	.16−.25
Superficial injury, open wounds	158	.09	.06−.11
Burns	14	.10	.01−.19
Poisonings	16	.05	.02−.09
Other injury	51	.09	.06−.12

[1] CI = confidence interval

Table 6.2 presents the mean uni-conditional EQ-5D disability weights of the diseases. Chronic non-specific lung disease had lowest mean EQ-5D disability weight of 0.18. Arthrosis had highest mean disability weight (0.45).

Table 6.2 Mean uni-conditional disability weights of pre-existing diseases, obtained from EQ-5D data of injury patients at 9 months follow-up

Pre-existing disease	N	Mean	CI[1]
CNLD[2]	7	.18	.13−.22
Heart disease	11	.23	.18−.28
Diabetes	6	.35	.15−.56
Backache	8	.31	.07−.55
Arthrosis	30	.45	.31−.58
Rheumatoid arthritis	10	.32	.15−.49
Other disease	51	.31	.25−.37

[1] CI = confidence interval, [2] CNLD = Chronic non-specific lung disease

The mean observed and predicted comorbid disability weights, where the predicted disability weights were calculated with the three different adjustment approaches, are shown in Table 6.3.

Table 6.3 Mean observed and predicted comorbid disability weights (pre-existing diseases and injury)

		Predicted (CI[1])		
	Observed (CI[1])	Max limit approach	Additive approach	Multipl approach
CNLD[2] and injury	.24 (.16−.33)	.20 (.14−.25)	.32 (.23−.41)	.29 (.22−.37)
Heart disease and injury	.31 (.20−.42)	.25 (.20−.30)	.39 (.30−.49)	.35 (.28−.43)
Diabetes and injury	.36 (.13−.59)	.35 (.15−.56)	.54 (.28−.79)	.47 (.26−.68)
Backache and injury	.33 (.09−.57)	.32 (.09−.56)	.44 (.19−.69)	.40 (.18−.62)
Arthrosis and injury	.53 (.39−.67)	.47 (.34−.59)	.61(.48−.75)	.55 (.43−.66)
Rheumatoid arthritis and injury	.44 (.29−.59)	.33 (.16−.50)	.50 (.34−.65)	.45 (.30−.59)
Other disease and injury	.30 (.24−.37)	.32 (.26−.38)	.47 (.40−.54)	.42 (.36−.48)

[1] CI = confidence interval

[2] CNLD = Chronic non-specific lung disease

Figure 6.2 shows the relationship between the observed and predicted disability weights for the injury types and comorbid diseases. Pearson's correlation coefficient and ICC between the observed and predicted comorbid disability weights were significant ($p<0.001$) for all three approaches, and highest for the multiplicative approach (Pearson's $r = 0.79$, ICC $=0.853$).

As shown in table 6.4, in case of severe chronic disease, all three methods showed a strong association of the predicted and observed morbid disability weight ($p<0.001$). For mild to moderate chronic disease the association between predicted and observed disability weights was lower, though still significant ($p<0.05$).

Table 6.4. Pearson's correlation coefficient of observed comorbid disability weights and comorbid disability weights predicted with the three adjustment approaches, stratified by severity of the disease

Severity	Maximum limit approach	Additive approach	Multiplicative approach
Mild to moderate disease	R=.383, p=.028	R=.373, p=.032	R=.370, p=.034
Severe disease	R=.865, p<.001	R=.884, p<.001	R=.880, p<.001

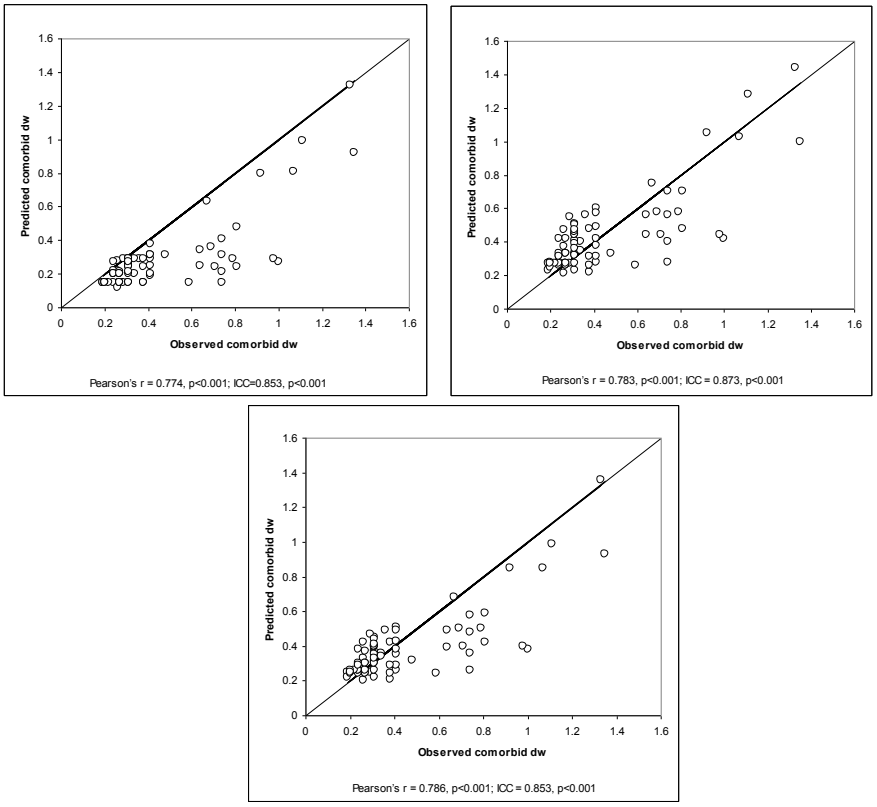

Figure 6.2 Observed and predicted disability weights for comorbid conditions (upper left: maximum likely approach; upper right: additive approach; lower middle: multiplicative approach)

Discussion

The results of this study showed that the EQ-5D disability weight of injury patients increases with the number of comorbid diseases. The three methods to account for comorbidity effects (i.e. the maximum limit, additive and multiplicative approach) provided reasonable results, especially in case of severe chronic disease.

The results of the current study are in line with the results of Flanagan et al. that tested the multiplicative approach using empirical utility data from the Canadian Community Health Survey [12]. They showed that observed and predicted utility was highly associated. It should, however, be noted that opposed to the current study, Flanagan et al. made no distinction between unrelated, indirectly related and directly related comorbidity and that only the multiplicative approach was tested.

Furthermore, Flanagan et al.[12] measured health loss using the Health Utilities Index (HUI), whereas in the current study the EQ-5D was used. Both are generic instruments, i.e. the functional health state of the patient and the utility weights derived from the

population are based on generic attributes and without regard to the underlying condition, disease specific key symptoms, prognosis or treatment. Evidence suggested that the sensitivity of the EQ-5D is low compared to HUI, implying that the EQ-5D does not measure disability where the HUI does [13]. Additionally, Polinder et al. showed that among injury patients the HUI is more sensitive for comorbid disease compared to the EQ-5D [14]. This might be caused by the crudeness of the levels, three in case of the EQ-5D compared to five or six levels of the HUI. Comparison of the visual analogue scale (VAS) scores of patients with and without comorbid disease might clarify this issue. The VAS valuation technique requires patients to score their health state on a vertical thermometer with endpoints 0 (worst possible health state) and
100 (best possible health state). Unfortunately, it was not possible to assess the effects of comorbid disease and injury with the VAS, because too many VAS scores were missing.

An important limitation that applies to the current study and the study of Flanagan et al.[12] is that utility scores were used to test the comorbidity adjustment approaches rather than the impact on the separate health domains of the multi-attribute utility instruments. This limitation may be overcome by a fourth adjustment approach that starts from the domain specific impact of a condition without comorbidity and compares this impact with the estimated domain impact of the comorbidity only. By selecting the maximum impact for each domain, a maximum limit profile is derived, for each domain the worst of both conditions. Subsequently, the total utility and disability weight of the maximum profile are convention-ally calculated. This approach can accommodate co-existing diseases which share affected do-mains, or the presence of two or more comorbid diseases. It does, however, require detailed descriptive data.

A second limitation that may have affected the results of the current study is that health loss of co-existing disease and temporary consequences of injury was measured using patient-reported EQ-5D data. As a result, adaptation might have affected the patient-reported EQ-5D data. The selected co-existing diseases were chronic and adaptation to their chronic health state might have caused patients to value their health state as less severe. This effect is especially found regarding chronically ill patients [15, 16]. The level of adaptation possibly differs between patients with mild to moderate and severe chronic disease and this may explain the differences found in association between predicted and observed disability weights in case of severe chronic disease compared to mild to moderate chronic disease.

A third limitation of the current study is that pre-injury utility scores of the injury patients were unavailable. Therefore, 9-month disability weight were used to calculate the comorbid disease with the three comorbidity adjustment approaches. We assumed that the 9-month disability weights of patients that were restored from their injury capture the health loss due to chronic, pre-existing disease without the effects of the injury and that the health loss caused by the pre-existing comorbid disease at 9 months was similar to the health loss at 2 months post-injury. However, the severity of the disease might have changed over time. Regarding rheumatic arthritis, for instance, symptoms may vary over

time. This may have resulted in either under- or overestimated the health loss due to the pre-existing disease at two months post-injury.
Furthermore, evidence suggests that EQ-5D utility scores of individuals may be influenced by medical as well as non-medical factors, such as age and sex, educational level, and marital status [17, 18]. In the current study, however, these non-medical factors that may have affect the reported EQ-5D health states reported by the injury patients were not taken into account.

After its use in the groundbreaking Global Burden of Disease studies, the DALY gained wide adherence, aiding national and international decision making processes regarding resource allocation and prevention [1, 2]. In many countries the DALY has been adopted as the standard outcome measure in public health and the primary outcome for priority setting in health care and prevention. However, by assuming a single disease hypothesis, the standard application of the DALY metric fails to consider the fact that more than one condition may exist simultaneously in a patients [19]. Apart from the difficulty to distinguish primary and secondary diagnosis, ignoring comorbid disease in burden of disease estimates restricts the use of the DALY in multi-morbid populations, such as the elderly in high income countries. Gains from care may easily be overestimated if a substantial number of patients suffer from additional conditions and in case disorders are more often held to be secondary than primary, their significance in burden of disease studies and the benefit of eradication rather is under-estimated. This has implications for priority setting and prevention and may lead to wrongful policy recommendations.
The results of this study showed that the performance available comorbidity adjustments methods is reasonable and that they may be applied in burden of disease studies. To improve current approaches to adjust for comorbidity we recommend more research on the effects of comorbidity on health-related quality of life. In this study the effects of temporary injuries and chronic disease have been investigated. Whether the effect of multiple chronic diseases on health-related quality of life and the performance of the available adjustment approach is the same remains to be investigated.
Next to the performance of currently available approaches to adjust for comorbidity, other adjustment approaches should be developed. The currently available methods encroach upon the disability weight, yet other methods, such as a domain specific should be explored.
Lastly, it is vital to collect data on comorbidity and to enhance registry of secondary diagnosis in vital health registries.

References

1. Murray, C.J.L., A.D. Lopez, and C.D. Mathers, Summary measures of population health: concepts, ethics, measurement and applications. 2002, Geneva: World Health Organization.
2. Field, M.J. and M.R. Gold, Summarising population health: Directions for the development and application of population health metrics. . 1998, Washington D.C.: Institute of Medicine: National Academy Press.
3. Bonsel, G.J., M.F. Janssen, and E. Birnie, Mild diseases & Ailments Study. Ziektelast-bepaling van 27 vermoedelijk 'lichte' ziekten ten behoeve van rationele beperking van het geneesmiddelenpakket. 2004, University of Amsterdam, Academic Medical Center: Amsterdam.
4. Mathers, C.D., K.M. Iburg, and S. Begg, Adjusting for dependent comorbidity in the calculation of healthy life expectancy. Popul Health Metr, 2006. 4: p. 4.
5. van Baal, P.H., et al., Disability weights for comorbidity and their influence on health-adjusted life expectancy. Popul Health Metr, 2006. 4: p. 1.
6. Polinder, S., et al., Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma, 2007. 62(1): p. 133-41.
7. Brooks, R., EuroQol: the current state of play. Health Policy, 1996. 37(1): p. 53-72.
8. Rabin, R. and F. de Charro, EQ-5D: a measure of health status from the EuroQol Group. Ann Med, 2001. 33(5): p. 337-43.
9. Dolan, P., Modeling valuations for EuroQol health states. Med Care, 1997. 35(11): p. 1095-108.
10. CBS, Permanent onderzoek leefsituatie. . 2003, Statistics Netherlands (CBS): Voorburg/Heerlen.
11. Mosby, Mosby's Medical Dictionary,. 8th ed. 2009: Elsevier Health Sciences.
12. Flanagan, W., et al., Deriving utility scores for co-morbid conditions: a test of the multiplicative model for combining individual condition scores. Popul Health Metr, 2006. 4: p. 13.
13. Bas Janssen, M.F., E. Birnie, and G.J. Bonsel, Evaluating the discriminatory power of EQ-5D, HUI2 and HUI3 in a US general population survey using Shannon's indices. Qual Life Res, 2007. 16(5): p. 895-904.
14. Polinder, S., et al., The measurement of long-term health related quality of life after injury; comparison of EQ-5D and the Health Utilities Index. Accepted for publication in Injury Prevention.
15. Adang, E.M., et al., Quality-of-life ratings in patients with chronic illnesses. Jama, 1997. 277(13): p. 1038.
16. Postulart, D. and E.M. Adang, Response shift and adaptation in chronically ill patients. Med Decis Making, 2000. 20(2): p. 186-93.
17. Kind, P., G. Hardman, and S. Macran, UK population norms for EQ-5D. 1999, Centre of Health Economics, University of York: York.
18. Hoeymans, N., H. van Lindert, and G.P. Westert, The health status of the Dutch population as assessed by the EQ-6D. Qual Life Res, 2005. 14(3): p. 655-63.
19. Anand, S. and K. Hanson, Disability-adjusted life years: a critical review. J Health Econ, 1997. 16(6): p. 685-702.

Part III

Long-term sequelae

7

Disease burden of post-infectious irritable bowel syndrome in The Netherlands

Juanita A. Haagsma, Peter D. Siersema, Niek J. De Wit, Arie H. Havelaar

Epidemiology and Infection 2010; 12: 1-7.

Abstract

Post-infectious irritable bowel syndrome (PI-IBS) has been established as a sequel of infectious intestinal disease (IID). The aim of this study was to estimate the burden of PI-IBS caused by the pathogens *Campylobacter*, *Salmonella* and *Shigella*, and to compare this with other outcomes associated with these pathogens. The attributable risk of PI-IBS due to bacterial pathogens was calculated and linked to national data on gastroenteritis incidence and measures for severity and duration of illness in order to estimate the burden of PI-IBS. One year post-infection, IBS developed in 9% of patients with bacterial IID. The burden of PI-IBS adds over 2300 disability adjusted life years to the total annual disease burden for the selected pathogens. PI-IBS is a frequent sequel of IID, resulting in a considerable disease burden compared to other outcomes. If this relationship is not considered, this will result in an underestimation of the disease burden of IID.

Introduction

Over the past decade, burden of disease estimates have become of increasing significance in allocating medical resources, in targeting interventions, and for monitoring possible effect. These studies can use monetary units to quantify the burden of disease, so-called cost-of-illness studies, or can assess the disease burden in a population by utilizing the Disability Adjusted Life Years (DALY) metric [1].

The DALY integrates the effect of mortality as well as disease and disability, which allows them to be considered at the same time. This simplifies comparisons between distinct disease outcomes and subgroups of a population. A prerequisite for these comparisons, either in monetary units or in DALYs, is that the burden of disease comprises a complete assessment of the health effect due to a certain pathogen. This means for infectious intestinal disease (IID) that, apart from the direct consequence of gastroenteritis, all sequelae should be con-sidered. In previous studies on the burden of IID [2-4], we included outcomes, such as Guillain-Barré syndrome, reactive arthritis, inflammatory bowel disease and haemolytic uraemic syndrome, leading to chronic renal failure to complement the burden of acute gastroenteritis.

Previous studies have shown that the prevalence of irritable bowel syndrome (IBS) in Western populations is between 10 to 20% [5, 6]. There is growing evidence of persisting gastrointestinal symptoms after bacterial IID, and these patients will, at least partly, meet the diagnostic criteria for (post-infectious) irritable bowel syndrome (PI-IBS) [7-11]. Although not life-threatening or severely debilitating, IBS symptoms may have a severe effect on daily life, affecting work, school and social life, and decreasing health-related quality of life over long periods [12, 13]. These periods of decreased health-related quality of life on the one hand and its high prevalence in society on the other might add up to a substantial burden of disease, which is currently not considered in estimates of the burden of IID.

Nonetheless, when PI-IBS is considered in such studies, the background IBS population prevalence and the percentage attributable to IID should be indisputably clear. For this purpose, the attributable risk (AR) has been established as the preferred measure of association.

In the current study, we first aimed to assess the AR of developing PI-IBS for bacterial pathogens. Second, we estimated the disease burden in DALYs of PI-IBS due to a particular group of pathogens, i.e. (thermophilic) *Campylobacter* spp., *Salmonella* spp., and *Shigella* spp. in The Netherlands. These pathogens were chosen because they are the most frequent causes of bacterial IID and have been demonstrated to be responsible for the largest burden of IID [2]. Third, we compared this burden with other outcomes associated with these pathogens.

Methods

Literature review

We based our study on the recently published meta-analysis by Thabane et al. [11]. From this paper we selected four studies (Table 7.1) in which:

- IID patients were considered who had no previous history of IBS or other bowel disorders;
- Appropriate control groups were included;
- A bacterial aetiology was confirmed by stool culture from at least a proportion of patients with IID.

In an additional search in the recent literature, we identified one study [14] that met the same selection criteria. This study presented follow-up data of a previously reported cohort [15, 16], and we therefore only used the most recent of these two studies in our evaluation. If available, we used data after a follow-up period of 1 year, or as close to this time point as possible. Studies involving patients with traveller's diarrhoea were excluded because it has been suggested that these are predominantly caused by enterotoxigenic *E. coli*, which has a milder course and a lower risk of developing PI-IBS [17].

AR

The AR was calculated as the additional incidence rate (IR) of exposed (IID) cases compared to unexposed (non-diseased) controls:

$$AR = IR_{exposed} - IR_{unexposed}.$$

The uncertainty in the AR was modelled by employing a Bayesian approach [18]. Both $IR_{exposed}$ and $IR_{unexposed}$ cases were considered as binomial fractions in which the uncertainty was modelled as a Beta (s, f) distribution, in which s is the number of persons in the exposed or unexposed group who developed IBS and f is the number of persons who did not develop IBS. With this approach, the 95% confidence interval (CI) is similar to that estimated by frequentist statistics, with the advantage that now a full uncertainty distribution can be simulated.

DALYs

The incidence of gastrointestinal infectious disease due to *Campylobacter*, *Salmonella* and *Shigella* was derived from national incidence data on foodborne disease, based on a method reported by Havelaar et al. [19]. The estimates were updated to the year 2006 (for details, see [2]). These incidence data relate to all cases in the population and are based on data from Sensor, a Dutch community-based cohort study [20]. By combining the incidence of gastrointestinal infectious disease associated with the selected pathogens and

a weighted mean AR, the incidence of PI-IBS was calculated. The estimated incidence of PI-IBS was then used to calculate the burden of PI-IBS in DALYs. Uncertainty in incidence and DALY estimates was quantified by Monte Carlo simulation.
The DALY aggregates mortality, expressed in years of life lost (YLL) and morbidity, expressed in years lived with disability, according to the following calculation:

DALY = YLL + YLD,

where YLL represents the time lost due to premature mortality. For PI-IBS the mortality component of the DALY is not taken into account as IBS is not associated with increased mortality. YLD represents the time in good health lost while living with a disease or disability and is calculated with the following formula:

$$YLD = \Sigma n_l * t_l * dw_l$$

where n_l is the number of cases with health outcome l, t_l the duration of the health outcome and dw_l the disability weight assigned to the health outcome. The disability weight is a value ranging from 0 to 1 that is assigned to living with a medical condition. This value reflects the effect of a specific health condition on health-related quality of life and is commonly based on the preferences of an expert or lay panel. We adopted the IBS disability weight from the Mild Diseases and Ailments Study, a Dutch study that generated disability weights for 52 health conditions from panels of medical experts and lay men recruited from the general population. In this study a renewed methodology was applied that focused especially on obtaining and improving disability weights for functional losses of a temporary and complex nature [21]. This can also be applied to IBS symptoms, because they tend to subside and exacerbate with periods of partial remission. To align with the societal point of view of the DALY, we used the IBS disability weight derived from the population panel (n=105) which had a value of 0.042. For the duration of PI-IBS, several long-term follow-up studies were assessed [15, 16, 22-25]. Based on these studies, we assumed for PI-IBS provoked by bacteria an average duration of 5 years [14].

Results

Table 7.1 presents an overview of the results of the five case-control studies that analyzed the incidence of PI-IBS following IID. In these studies, the index group was a cohort of gastroenteritis patients with confirmed bacterial aetiology. Controls in these studies were either volunteers, patients from the same general practice or siblings or spouses of the patient. The PI-IBS incidence rate in these five studies ranged from 4% [26] up to 17% [27]. The calculated AR is also shown in table 7.1. These data do not suggest that PI-IBS is different for various bacterial pathogens (χ^2 test for homogeneity, p=0.95 [28]). We

therefore pooled the studies considering patients with bacterial aetiology, weighted by the number of cases that experienced an episode of IID. This resulted in a mean weighted AR for bacterial infections (10-12 months post-IID) of 8.8% (90% CI: 7.2-10.4%, figure 7.1).

Table 7.1 Population attributable risk of developing IBS following gastroenteritis, per study

Study	Aetiology	Follow-up (months)	Incidence of IBS		AR (%)
			Cases (%)	Controls (%)	
Rodriguez et al. [26]	Bacterial	12	14/318 (4.4)	2027/584308 (0.3)	4.1
Parry et al. [27]	*Campylobacte, Salmonella*	6	18/108 (16.7)	4/206 (1.9)	14.8
Wang et al. [29]	*Shigella*	10	24/235 (10.2)	2/243 (0.8)	9.4
Mearin et al. [30]	*Salmonella*	12	31/266 (13.2)	5/333 (1.5)	11.7
Jung et al. [14]	*Shigella*	12	12/87 (13.8)	1/88 (1.1)	12.7

IBS, Irritable bowel syndrome; AR, Attributable risk.

In The Netherlands, about 124 000 cases of IID due to the three pathogens were considered in 2006 (Table 7.2). Most of these IID cases were caused by *Campylobacter* (78 000), and *Salmonella* (43 000). About 27 000 GP visits were registered due to these three pathogens.

Table 7.2 Incidence of infectious intestinal disease due to three pathogens in The Netherlands (population 16.3 million), 2006

Pathogen	Total cases	GP visits	Reported cases
Salmonella	42,000 (8,800-110.000)*	7,300 (4,300-11,000)	1,700
Campylobacter	78,000 (28.000-170.000)	19,000 (10.000-32,000)	3,400
Shigella	4,400 (600-11,000)	700 (100-1,800)	240

GP, General Practitioner

* mean (5-95 percentile)

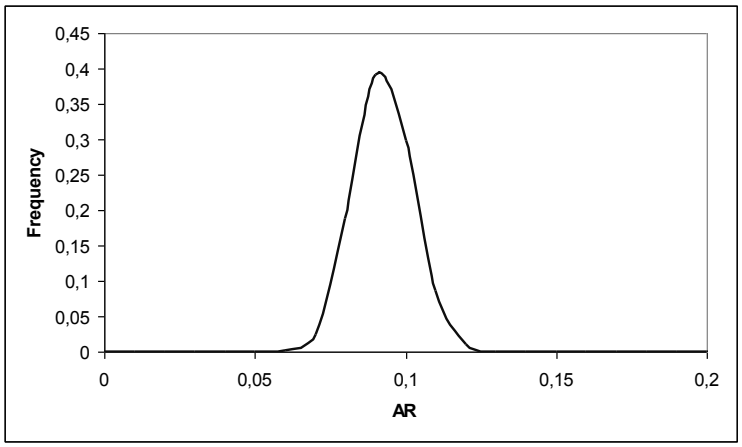

Figure 7.1 Uncertainty distribution of the attributable risk (AR) of post-infectious irritable bowel syndrome in infectious intestinal disease patients with bacterial aetiology

The three pathogens were estimated to be associated with about 11 000 new cases of PI-IBS per year (table 7.3), in this way contributing 55 000 cases to the overall prevalence of IBS in The Netherlands; which is estimated to be 330 000 formally diagnosed patients [5].

Table 7.3 Incidence and prevalence of post-infectious irritable bowel syndrome due to three pathogens in The Netherlands, 2006

Pathogen	Incidence	Prevalence[$]
Salmonella	3700 (750-9,300)[*]	19000 (3800-46000)
Campylobacter	6900 (2400-15,000)	34000 (12000-74000)
Shigella	370 (50-960)	1800 (260-4800)

[$] Incidence x 5 years

[*] mean (5-95 percentile)

The disease burden of PI-IBS is shown in table 7.4. In the year 2006, an estimated 2302 DALYs were lost due to PI-IBS following IID induced by the three selected pathogens, with 63% of DALYs due to *Campylobacter*, 34% to *Salmonella* and 3% to *Shigella*. Including PI-IBS in this calculation increased the burden of disease (figure 7.2) for *Salmonella* by 86% (from 905 to 1686 DALY), for *Campylobacter* by 92% (from 1564 to 3008 DALY) and for *Shigella* by 151% (from 51 to 128 DALY).

Table 7.4 Disease burden of post-infectious irritable bowel syndrome due to four pathogens in The Netherlands, 2006

Pathogen	Incidence	Duration (years)	Disability weight	Disease burden (YLD)
Salmonella	3,700 (750-9,300)*	5	0.042	780 (160-2,000) *
Campylobacter	6,900 (2,400-15,000)	5	0.042	1,400 (510-3,100)
Shigella	370 (50-960)	5	0.042	80 (10-200)

YLD, Years Lived with Disability, * mean (5-95 percentile)

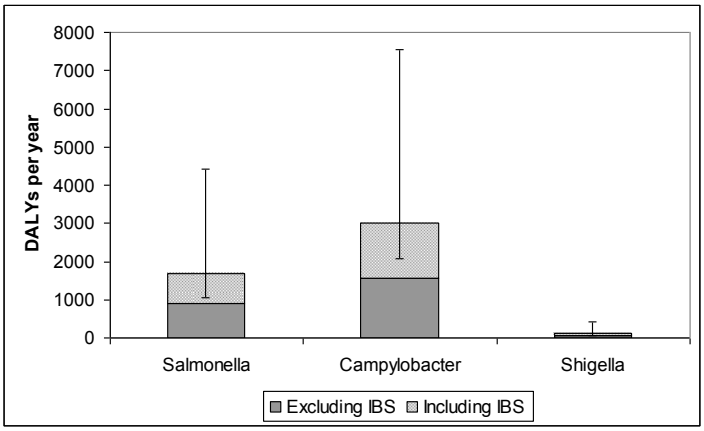

Figure 7.2 Disease burden in disability adjusted life years (DALYs) per year due to three pathogens in The Netherlands, 2006, including and excluding post-infectious irritable bowel syndrome (IBS)**.** Error bars express the 5th and 95th percentiles resulting from Monte Carlo simulation.

Discussion

Based on recent findings from the literature, we estimated that approximately 9% of patients with bacterial IID PI-IBS will develop 1 year after infection [11]. We also estimated that approximately 55 000 prevalent cases of PI-IBS in the Dutch population will actually result from an infection with *Salmonella*, *Campylobacter* or *Shigella*. This is about 17% of the total prevalence of formally diagnosed PI-IBS. Others [31] have estimated a similar range of 6-17%.

Our study used data from case-control studies to assess the incidence and disease burden of PI-IBS. Over the past few years, the evidence that an episode of IID is a risk factor for developing IBS has increasingly been recognized, reinforced by a growing understanding of the pathophysiology of IBS [32-34]. In brief, mucosal injury and inflammation are assumed to increase the number of enterochromaffin cells, thus increasing the release of serotonin (5-hydroxytryptamine, 5-HT). 5-HT affects gastrointestinal motility, enterocyte

secretion and visceral sensation. Increased levels of pro-inflammatory cytokines and lymphocyte counts in intestinal epithelium and lamina propria have indeed been documented in PI-IBS patients [7].

In a recent follow-up study, 3 years after a case-control study on laboratory-confirmed cases of *Campylobacter* and *Salmonella*, an increased risk of IBS could not be confirmed [35]. In contrast, the results of this study suggested that patients with IBS or with another chronic intestinal condition were more susceptible to IID. The studies identified in our literature review excluded cases with pre-existing symptoms; hence, we only included cases in which the IID episode triggered a new case of IBS. Patients with already existing IBS or another intestinal illness were not included in our study.

In the current study, only limited information was available on some other factors involved in PI-IBS, and this may have affected our results in the following way. First, the data used to estimate the incidence of PI-IBS were retrieved from the existing literature. Several case-control studies were used and the time interval chosen for follow-up of exposed and non-exposed participants varied from 3 months [15] to 5 years [14]. Overall, we found that the incidence of PI-IBS among exposed participants was higher with a shorter follow up period [22, 23, 30]. In order to present a pooled estimate we chose a follow-up period of 10-12 months. Using shorter follow-up periods would have resulted in an increased AR, and, subsequently, a higher estimated incidence of PI-IBS. As these additional cases had a disease duration of <1 year, they would have added little value to the total burden of PI-IBS.

Second, in this study we did not consider patient characteristics (such as gender, smoking or psychosocial co-morbidity) as risk factors for developing PI-IBS. Similarly, we did not take into account the details of the gastroenteritis episode, such as duration or aetiology of the infection [10, 22, 25, 27, 36]. With regard to duration of gastroenteritis, previous studies revealed that the longer period patients have symptoms due to gastroenteritis or the larger effect these symptoms have, the higher the risk of having PI-IBS [36]. We were not able to adjust for IID duration or severity, due to a lack of comparison.

Third, the disability weight is based on the preferences of a population panel, which is in line with the societal perspective of the DALY metric. Nevertheless, the IBS disability weight is not able to capture the heterogeneous group of clinical symptoms that are experienced by patients. Several studies have shown that there is a wide variation in symptoms between patients [37]. As a result, actual health states of patients with (PI-)IBS may differ considerably from the health state descriptions valued by the population panel. Moreover, the value of the disability weight seems to be rather low. Health-related quality of life studies performed among patients with IBS have shown that the loss of health-related quality of life was indeed considerable [13, 38]. This is not reflected by the value of the disability weights as applied in this study.

A case-control study by Marshall et al. has suggested that PI-IBS may also result from presumed viral IID [23]. As many as 20% (95% CI: 11-29)of PI-IBS cases developed in IID patients, although this risk was only statistically significant after 3 months. This shorter duration of PI-IBS due to viral IID is in line with observations of Spiller and

Garsed [31], who noted that recent studies have demonstrated that there is very little mucosal destruction in viral IID and, therefore, that the mucosal structure is likely to rapidly return to normal. When the shorter duration is taken into account, PI-IBS due to norovirus would still accumulate to a prevalence estimate of 32 000 cases in the population and a disease burden of 1300 YLD, which is comparable to the disease burden of PI-IBS due to *Campylobacter* infection. This implies that including PI-IBS in disease burden estimates of the selected bacterial pathogens might still underestimate the burden of disease of foodborne disease. Further follow-up studies are needed that provide more information on the risk of IBS after viral IID.

The results from the current study show that PI-IBS is a frequent sequel of IID and that the disease burden of PI-IBS is considerable compared to other outcomes of IID. The resulting annual non-fatal burden of disease of PI-IBS is over 2300 YLD, almost doubling the total non-fatal burden of disease for the selected pathogens. Ignoring PI-IBS in burden of disease studies results in an underestimation of the size of the burden of these types of gastrointestinal diseases. In order to satisfy the aspirations of burden of disease studies, i.e. identifying priorities in medical resource allocation, targeting of interventions and monitoring possible effect, all sequelae of IID should be included, given that the evidence on the association is sufficiently conclusive. There is definitely more research required on the duration of PI-IBS and related fluctuations in health-related quality of life, as well as on the association of viral gastroenteritis and PI-IBS.

References

1. Murray, C.J.L. and A.D. Lopez, The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020. 1996, Cambridge: Harvard University Press.
2. Kemmeren, J.M., et al., Priority setting of foodborne pathogens. Disease burden and costs of selected enteric pathogens. 2006, National Institute of Public Health and the Environment: Bilthoven.
3. Cressey, P. and R. Lake, Risk ranking: estimates of the burden of foodborne disease for New Zealand. 2007, ESR: Christchurch.
4. Havelaar, A.H., et al., Health burden in the Netherlands due to infection with thermophilic Campylobacter spp. Epidemiol Infect, 2000. 125(3): p. 505-22.
5. Hungin, A.P., et al., The prevalence, patterns and impact of irritable bowel syndrome: an international survey of 40,000 subjects. Aliment Pharmacol Ther, 2003. 17(5): p. 643-50.
6. Mertz, H.R., Irritable bowel syndrome. N Engl J Med, 2003. 349(22): p. 2136-46.
7. DuPont, A.W., Postinfectious irritable bowel syndrome. Clin Infect Dis, 2008. 46(4): p. 594-9.
8. Halvorson, H.A., C.D. Schlett, and M.S. Riddle, Postinfectious irritable bowel syndrome--a meta-analysis. Am J Gastroenterol, 2006. 101(8): p. 1894-9; quiz 1942.
9. Watson, W.L., J. Ozanne-Smith, and J. Richardson, Retrospective baseline measurement of self-reported health status and health-related quality of life versus population norms in the evaluation of post-injury losses. Inj Prev, 2007. 13(1): p. 45-50.
10. Spiller, R.C., Role of infection in irritable bowel syndrome. J Gastroenterol, 2007. 42 Suppl 17: p. 41-7.

11. Thabane, M., D.T. Kottachchi, and J.K. Marshall, Systematic review and meta-analysis: The incidence and prognosis of post-infectious irritable bowel syndrome. Aliment Pharmacol Ther, 2007. 26(4): p. 535-44.
12. Glia, A. and G. Lindberg, Quality of life in patients with different types of functional constipation. Scand J Gastroenterol, 1997. 32(11): p. 1083-9.
13. Pare, P., et al., Health-related quality of life, work productivity, and health care resource utilization of subjects with irritable bowel syndrome: baseline results from LOGIC (Longitudinal Outcomes Study of Gastrointestinal Symptoms in Canada), a naturalistic study. Clin Ther, 2006. 28(10): p. 1726-35; discussion 1710-1.
14. Jung, I.S., et al., The Clinical Course of Postinfectious Irritable Bowel Syndrome: A Five-year Follow-up Study. J Clin Gastroenterol, 2009.
15. Ji, S., et al., Post-infectious irritable bowel syndrome in patients with Shigella infection. J Gastroenterol Hepatol, 2005. 20(3): p. 381-6.
16. Matsuoka, Y., et al., Incidence and prediction of psychiatric morbidity after a motor vehicle accident in Japan: the Tachikawa Cohort of Motor Vehicle Accident Study. Crit Care Med, 2008. 36(1): p. 74-80.
17. DuPont, H.L., Systematic review: prevention of travellers' diarrhoea. Aliment Pharmacol Ther, 2008. 27(9): p. 741-51.
18. Vose, D., Risk Analysis, A Quantative Guide 2000, Chichester: John Wiley and Sons.
19. Havelaar, A.H., et al., Disease burden in The Netherlands due to infections with Shiga toxin-producing Escherichia coli O157. Epidemiol Infect, 2004. 132(3): p. 467-84.
20. de Wit, M.A., et al., Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol, 2001. 154(7): p. 666-74.
21. Janssen, M.F., E. Birnie, and G. Bonsel, Feasibility and reliability of the annual profile method for deriving QALYs for short-term health conditions. Med Decis Making, 2008. 28(4): p. 500-10.
22. Gwee, K.A., et al., Psychometric scores and persistence of irritable bowel after infectious diarrhoea. Lancet, 1996. 347(8995): p. 150-3.
23. Marshall, J.K., et al., Postinfectious irritable bowel syndrome after a food-borne outbreak of acute gastroenteritis attributed to a viral pathogen. Clin Gastroenterol Hepatol, 2007. 5(4): p. 457-60.
24. McKeown, E.S., et al., Postinfectious irritable bowel syndrome may occur after non-gastrointestinal and intestinal infection. Neurogastroenterol Motil, 2006. 18(9): p. 839-43.
25. Neal, L.A., et al., Convergent validity of measures of post-traumatic stress disorder in a mixed military and civilian population. J Trauma Stress, 1994. 7(3): p. 447-55.
26. Rodriguez, L.A. and A. Ruigomez, Increased risk of irritable bowel syndrome after bacterial gastroenteritis: cohort study. Bmj, 1999. 318(7183): p. 565-6.
27. Parry, S.D., et al., Does bacterial gastroenteritis predispose people to functional gastrointestinal disorders? A prospective, community-based, case-control study. Am J Gastroenterol, 2003. 98(9): p. 1970-5.
28. Fisher, L.D. and G. van Belle, Biostatistics: A Methodology for the Health Sciences. 1993, New York: John Wiley and Sons.
29. Wang, L.H., X.C. Fang, and G.Z. Pan, Bacillary dysentery as a causative factor of irritable bowel syndrome and its pathogenesis. Gut, 2004. 53(8): p. 1096-101.
30. Mearin, F., et al., Dyspepsia and irritable bowel syndrome after a Salmonella gastroenteritis outbreak: one-year follow-up cohort study. Gastroenterology, 2005. 129(1): p. 98-104.
31. Spiller, R. and K. Garsed, Postinfectious irritable bowel syndrome. Gastroenterology, 2009. 136(6): p. 1979-88.
32. Crowell, M.D., et al., Enterochromaffin cells and 5-HT signaling in the pathophysiology of disorders of gastrointestinal function. Curr Opin Investig Drugs, 2004. 5(1): p. 55-60.
33. Dunlop, S.P., et al., Abnormalities of 5-hydroxytryptamine metabolism in irritable bowel syndrome. Clin Gastroenterol Hepatol, 2005. 3(4): p. 349-57.

34. Houghton, L.A., J.M. Foster, and P.J. Whorwell, Alosetron, a 5-HT3 receptor antagonist, delays colonic transit in patients with irritable bowel syndrome and healthy volunteers. Aliment Pharmacol Ther, 2000. 14(6): p. 775-82.
35. Doorduyn, Y., et al., Novel insight in the association between salmonellosis or campylobacteriosis and chronic illness, and the role of host genetics in susceptibility to these diseases. Epidemiol Infect, 2008. 136(9): p. 1225-34.
36. Cuomo, R., M.F. Savarese, and R. Gargano, Almost all irritable bowel syndromes are post-infectious and respond to probiotics: consensus issues. Dig Dis, 2007. 25(3): p. 241-4.
37. Bijkerk, C.J., et al., Randomized patients in IBS research had different disease characteristics compared to eligible and recruited patients. J Clin Epidemiol, 2008. 61(11): p. 1176-81.
38. Brazier, J., et al., A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ, 2004. 13(9): p. 873-84.

8

Posttraumatic stress symptoms and health-related quality of life two years after injury treated at the Emergency Department

Juanita A. Haagsma, Suzanne Polinder, Miranda Olff, Hidde Toet, Gouke J. Bonsel, Ed F. van Beeck

Submitted

Abstract

Introduction Among injury victims relatively high prevalence rates of posttraumatic stress disorder (PTSD) have been found. However, previous studies that investigated prevalence rates of PTSD were mainly conducted in clinical patient populations and were therefore restricted to accidents and injuries at the higher end of the severity spectrum. To identify opportunities for prevention and early treatment, the prevalence and predictors of PTSD in a comprehensive population of injury patients of all severity levels and external causes need to be investigated. This study assessed the prevalence rate and predictors of PTSD symptoms in a comprehensive ED-treated population of injury patients two years post-injury. Also, the association between PTSD symptoms and health related quality of life (HRQoL) was examined .

Methods We conducted a self-assessment survey which included items regarding demographics of the patient, accident type, sustained injuries, EQ-5D and Health Utilities Index (HUI) to measure functional outcome and HRQoL, and the Impact of Event Scale (IES) to measure PTSD symptoms. An IES-score ≥ 35 was used as indication for the presence of PTSD. The survey was completed by 1,781 injury patients two years after they were treated at the ED, followed by either hospital admission or direct discharge to the home environment.

Results In this sample of injury patients, 4% reported symptoms indicative of PTSD (IES ≥ 35) two years after the injury. Independent predictors of this disorder were female gender, occupational injury, comorbid disease, hospitalization and severity level of the sustained injuries. Symptoms indicative of PTSD (IES ≥ 35) were associated with more problems on all EQ-5D and HUI3 domains of functional outcome and a considerable utility loss in both hospitalised (0.23-0.24) and non-hospitalized (0.32-0.33) patients. After adjusting for potential confounders, PTSD remained strongly associated with adverse HRQoL.

Conclusions Two years after their injury one out of 25 patients treated at an ED suffers from posttraumatic stress symptoms indicative of PTSD, which are associated with a considerable decrease in HRQoL in both hospitalised and non-hospitalized patients. PTSD symptoms may therefore raise a major barrier for full recovery of injury patients of even minor levels of severity.

Introduction

Posttraumatic stress disorder (PTSD) may result from any event that involves an injury, or threatened or actual death. Regarding injury victims PTSD prevalence rates up to 39% have been found one to four months after the injury [1]. At long-term follow-up (>1 year) PTSD prevalence rates from 5% [2] to 32% [3] have been reported. A substantial share of studies that investigated prevalence rates and predictors of PTSD following injury addressed certain injury subgroups, such as victims of motor vehicle accidents [4-7], burn victims [8-10] or patients who required admission to hospital or the Intensive Care Unit [3, 11-15]. Those previous studies were mainly conducted in clinical patient populations and were therefore restricted to accidents and injuries at the higher end of the severity spectrum. One prior publication suggested that PTSD is relatively common among injury patients treated at the Emergency Department [16], yet the respondents included in this study were admitted to the trauma service for at least 24 hours. Injuries cover a variety of causes and consequences and to identify opportunities for prevention and early treatment, the prevalence and predictors of PTSD in a comprehensive population of injury patients of all severity levels and external causes needs to be investigated.

PTSD is associated with functional impairments and decreased health-related quality of life (HRQoL) [17, 18]. In one of the scarce studies addressing the latter, Holbrook et al. [19] showed that in a subgroup of injury patients admitted to a trauma centre PTSD has a substantial impact on health-related quality of life. Similar results were found among adolescents and children [20, 21]. However, these studies were again restricted to victims at the higher end of the severity spectrum and the association between PTSD and health-related quality of life among a comprehensive population of injury patients has yet to be studied.

The objectives of this study were:

1) to assess the prevalence rate and predictors of posttraumatic stress symptoms indicative of PTSD two years post-injury in a population of injury patients treated at the ED, followed by either hospital admission or direct discharge to the home environment.

2) to assess the association between posttraumatic stress symptoms indicative of PTSD and HRQoL among this comprehensive injury population.

Methods

Study design

A patient-follow-up study, which was previously published [22], was conducted among a population-based sample of injury patients of all severity levels. This study followed injury patients aged 15 years and older who attended the ED of the Dutch Injury Surveillance System (a representative continuous registry of intentional and unintentional injuries of 17

hospitals in the Netherlands). Surveys were conducted at 2 months, 5 months, 9 months and two years after initial treatment.

Study population

Between 8 October 2001 and 31 December 2002 a sample was selected of 8,564 patients aged 15 years and older who attended the ED of the Dutch Injury Surveillance System [22]. The patients were treated at the ED, followed by either hospital admission or direct discharge to the home environment. The sample of patients consisted of victims of traffic, home and leisure, occupational and sport accidents. The sustained injuries varied from minor to severe injury, single and multiple injury and hospitalized and non-hospitalized patients. The sample of patients was stratified, over sampling patients who were hospitalised. Each injury patient of the selected sample received a postal questionnaire 2½ months after the injury and 3,167 (37%) responded. The first questionnaire was made anonymous for privacy reasons. At 5, 9 and 24 months a follow-up questionnaire was sent to patients that responded to the preceding questionnaire. For these questionnaires the patients needed to give permissions by an informed consent form. The 5, 9 and 24 months follow-up questionnaire were completed by respectively 2,384 , 2,295 and 1,781 patients. The present study used a sample of 1,781 respondents (i.e. 21% of the original sample) on the two year post-trauma survey, which assessed both posttraumatic stress symptoms and HRQoL [22]. To adjust the data for non-response, a non-response analysis was conducted [22]. Multivariate logistic regression analysis was used to examine if variables age, sex, type of injury, external cause of the injury, hospitalization and length of stay, health status and ambulance transport were possible determinants of non-response. The significant variables were used to adjust for response bias by inverse probability weighting [23]. Additionally, the data were adjusted for stratification of the sample of ED patients [22].

Questionnaire

The follow-up questionnaire included items regarding demographics of the patient, accident category, type of injury, health care use and the Impact of Event Scale (IES), which was used to assess symptoms of posttraumatic stress indicative of PSTD [24]. The IES consists of 15 items, which measure intrusive re-experiences of the trauma and avoidance of trauma-related stimuli. By combining the 15 items the total IES-score, ranging from 0 through 75, can be calculated. Wohlfarth et al. showed that a cut-off score of 35 on the total IES-score produced a sensitivity of .89, and a specificity of .94 against the DSM-IV diagnostic criteria for PTSD as the gold standard [25]. Therefore, we assumed that an IES-score higher than 35 (IES≥35) represents symptoms of posttraumatic stress indicative of PTSD. The Dutch translation of the IES has been found to be valid and reliable [26].
Additionally, the questionnaire included items to measure functional outcome and HRQoL. HRQoL is an index of perceived functional outcome of an illness and disability

that is anchored between 0 (worst imaginable health state or death) and 1 (full health), thus allowing comparison between the health status of patients with distinct diseases. To measure HRQoL, multi-attribute utility instruments (MAUIs) such as the Health Utility Index (HUI) or the EQ-5D may be used [27, 28]. These instruments require the patient to report his or her health state with a standardised generic health state classification system, which is then converted into a health utility score using utility weights derived from the general population. Despite the similarities in obtaining the health utility score, there are important variations between the instruments regarding the health domains included in the health classification system and the methods applied to derive the utility weights [29]. As a result of these variations, the distinct instruments yield different utilities for similar health states. To overcome omissions in measuring HRQoL it is important to use several instruments that have complementary health domains [30].
Therefore, to measure functional outcome and HRQoL, the questionnaire included the EQ-5D and the HUI mark 3 (HUI3). With the EQ-5D classification system, respondents describe their health in three levels of severity on the health domains mobility, self-care, usual activities, pain/discomfort and anxiety/depression [31]. Subsequently, the weight of that health state is computed by a formula that firstly yields a partial weight score for each domain depending on the reported level and secondly adds the utility weights (also referred to as the 'tariff'), which are based on preference data of the general population of Canada [32].
The questionnaire included 19 items regarding the presence of one or more chronic disease(s) prior to the injury to assess comorbidity [33]. Comorbidity is defined as the presence of any coexisting medical conditions or disease processes additional to the injury that the injury patients sustained [34].

Data-analysis

For analysis of the data the Statistical Package for the Social Sciences version 14.0 was used (SPSS Inc, Chigaco, Ill). Missing IES values were imputed by substituting the median value of nearby 5 points [35]. Chi-square statistics (dichotomous variables) and Student t tests (continuous variables) were used to test for differences between injury patients with IES scores higher or lower than 35.
Univariate logistic regression and multivariate logistic regression analyses (enter method) were used to determine the predictive value of patient demographics, accident category and severity level of the sustained injuries with regard to posttraumatic stress symptoms indicative of PTSD (IES≥35) at two-year post-trauma. To dichotomize severity level, the injury diagnoses were categorized into two severity classes (mild versus moderate to severe) as previously tested by an international expert group [36]. The injury severity class moderate to severe comprises injuries such a skull-brain injury, fracture/dislocation of the vertebral column, fracture of pelvis and hip fracture. The injury severity class mild comprises injuries such as superficial injury, concussion and wrist fracture.

For the analysis of the association between IES≥35 and HRQoL, we selected participants that filled in both EQ-5D and HUI3. To test differences between participants with and without PTSD regarding their responses on each of the EQ-5D and HUI3 health domains, the non-parametric Wilcoxon-Mann-Whitney test was conducted. Differences regarding the mean EQ-5D and HUI3 summary scores were tested with a one-way ANOVA. P-values < 0.05 were considered to indicate statistical significance.
Stepwise multiple regression analyses (enter method) was applied to investigate the association between demographics (block 1), hospitalization and comorbidity (block 2) posttraumatic stress symptoms indicative of PTSD (IES≥35) (block 3) and HRQoL measured with the EQ-5D and HUI3.

Results

Study population

Regarding the respondents on the 24-month follow-up questionnaire, the average age was 44.5 years old and 46% were female. Over one half (54%) was injured due to home and leisure accidents. The sustained injuries of all the respondents consisted mostly of superficial injury/open wounds (51%) and upper extremity fractures (13%). After treatment at the ED, 9% of the respondents were admitted to hospital. Approximately one third (31%) had one or more pre-existing comorbid conditions.

Prevalence and predictors of posttraumatic stress symptoms indicative of PTSD

Table 8.1 shows the characteristics of the respondents, stratified by presence of posttraumatic stress symptoms indicative of PTSD (IES≥35). [25]. In our sample of ED-treated injury patients of all severity levels, 4.1% had an IES-score of 35 or higher. Age, female sex, comorbid disease, accident category and hospitalization were significantly associated with PTSD. Multivariate logistic regression analysis including demographic variables, accident category and severity level, indicated that female gender (OR 4.0; 95% CI, 3.8 to 4.3; p<0.001) is a strong independent predictor of PTSD two years after injury, having a larger effect than occupational injury (OR 2.2; 95% CI 1.8 to 2.7; p<0.001) and hospitalization (OR 2.2; 95% CI 2.0 to 2.4; p<0.001) (see table 8.2).

Table 8.1 Characteristics of injury patients, accident category and hospitalization status, stratified by presence of PTSD (IES-score≥35)

Characteristics	No PTSD IES-score<34[B] (n=1707; 95.9%)	PTSD IES-score≥35 [B] (n=73; 4.1%)	p
Patient demographics			
Age	45.9 (sd[b] 23.1)	35.2 (sd[b] 15.9)	<0.001
Female sex	45%	74%	<0.001
Comorbid disease	31%	44%	<0.001
Accident category			
Home and leisure	54%	61%	<0.001
Traffic	16%	19%	<0.001
Occupational	13%	17%	<0.001
Sport	17%	3%	<0.001
Hospitalization	8%	12%	<0.001

[a] an IES-score ≥35 signifies symptoms of posttraumatic stress indicative of PTSD

[b] sd = standard deviation

Table 8.2. Odds ratios (OR) and 95% confidence interval (CI) for the association of PTSD (IES-score≥35) with characteristics of the respondent/injury

	Not adjusted		Adjusted[a]		
Characteristics	OR	95% CI	OR	95% CI	p
Patient demographics					
Age	0.94	0.89 – 0.99	0.59	0.55 – 0.63	<0.001
Female sex	3.66	3.46 – 3.87	4.00	3.76 – 4.26	<0.001
Comorbid disease	2.16	2.05 – 2.27	1.78	1.69 – 1.88	<0.001
Accident category					
Home and leisure	1.47	1.40 – 1.55	1.11	0.91 – 1.36	0.283
Traffic	1.20	1.13 – 1.28	0.98	0.81 – 1.19	0.858
Occupational	1.20	1.12 – 1.29	2.23	1.83 – 2.73	<0.001
Sport	0.13	0.13 – 0.15	0.18	0.14 – 0.23	<0.001
Severity level of the injury					<0.001
Moderate to severe injury	1.20	1.04 – 1.38	1.55	1.32 – 1.80	<0.001
Hospitalization	1.75	1.63 – 1.88	2.20	2.03 – 2.39	<0.001

[a] adjusted for all included predictors in model using multivariate logistic regression analysis

Association of posttraumatic stress symptoms indicative of PTSD with HRQoL

With reference to the 1,781 respondents that completed the 24-month follow-up questionnaire, 1,585 (89%) filled in the EQ-5D and the HUI3 and 1,380 (77.5%) filled in the IES.

EQ-5D – Table 8.3 shows the responses on the EQ-5D of injury patients with IES scores higher or lower than 35. The calculated mean EQ-5D summary score for injury patients with IES scores≥35 was 0.56, whereas for injury patients with lower IES scores the mean EQ-5D summary score was 0.87 (t=112.0; p<0.001). Respondents with posttraumatic stress symptoms indicative of PTSD reported significantly more problems on all five EQ-5D health domains (p < 0.001). Differences in reported problems between patients with IES scores higher or lower than 35 were largest for EQ-5D health domains pain/discomfort (82% versus 28%) and anxiety/depression (53% versus 11%).

Table 8.3 Mean utility scores and percentage of reported problems on the EQ-5D and HUI3 health domains of the respondents without and with PTSD (IES-score≥35)

	IES-score<34[B] (n=1585)	IES-score≥35[B] (n=93)	p
EQ-5D			
Mean EQ-5D utility score	0.87 (sd[b] 0.15)	0.56 (sd[b] 0.26)	<0.001
Problems with mobility	15.4%	47.0%	<0.001
Problems with self-care	5.1%	18.9%	<0.001
Problems with usual activities	16.9%	53.2%	<0.001
Pain/discomfort	28.1%	82.3%	<0.001
Anxiety/depression	11.4%	53.9%	<0.001
HUI3			
Mean HUI3 utility score	0.83 (sd[a] 0.24)	0.51 (sd[a] 0.26)	<0.001
Problems with vision	54.1%	65.2%	<0.001
Problems with hearing	8.6%	1.6%	<0.001
Problems with speech	4.6%	28.7%	<0.001
Problems with ambulation	12.8%	20.4%	<0.001
Problems with dexterity	10.6%	23.0%	<0.001
Problems with emotion	32.5%	91.6%	<0.001
Problems with cognition	14.3%	52.6%	<0.001
Pain	38.4%	84.4%	<0.001

Mann-Whitney tests were performed for each domain

[a] an IES-score ≥35 signifies symptoms of posttraumatic stress indicative of PTSD

[b] sd = standard deviation

When the responses of hospitalized and non-hospitalized injury patients with IES≥35 and IES<35 are presented separately, again patients with symptoms indicative of PTSD (IES≥35) report significantly more problems on each of the EQ-5D health domains, resulting in a mean EQ-5D utility loss of 0.32 for non-hospitalized patients (t=112.2; p<0.001) and 0.23 for hospitalized patients (t=22.1; p<0.001). Compared to the injury patients without PTSD indications, injury patients with symptoms indicative of PTSD (IES≥35) at 24-months post-trauma also had significantly lower mean EQ-5D utility scores at the 2½-month (t=105.0, p<0.001), 5-month (t=100.1, p<0.001) and 9-month (t=38.1, p<0.001) follow-up, (see figure 8.1).

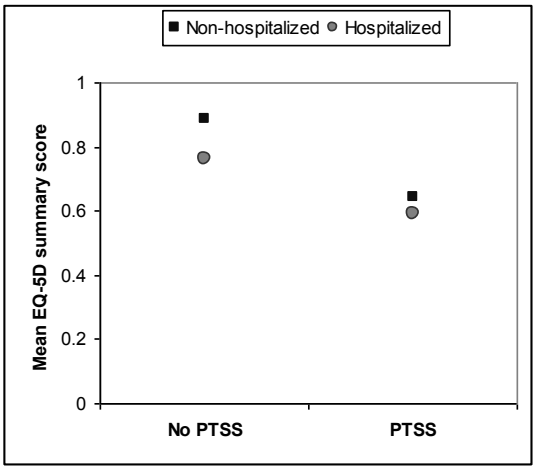

Figure 8.1 Mean EQ-5D utility scores of hospitalized and non-hospitalized patients with and without symptoms indicative of post-traumatic stress disorder (PTSS) at 24-month

HUI3 – Table 8.3 also shows the responses on the HUI3 domains reported by injury patients with IES-scores higher or lower than 35. For patients with IES≥35 the calculated mean HUI3 summary score was 0.51 and for patients with lower IES-scores 0.83 (t=81.1; p<0.001). Respondents with posttraumatic stress symptoms indicative of PTSD (IES≥35) reported significantly more problems on all HUI3 health domains, except hearing where a reverse association was found ($p < 0.001$). Differences in reported problems between patients with IES≥35 and IES<35 were largest for the HUI3 health domains emotion (92% versus 33%) and pain (84% versus 38%). Analysing the responses of non-hospitalized and hospitalized patients with and without PTSD indications (IES≥35) separately shows that hospitalized patients with symptoms indicative of PTSD (IES≥35) reported most problems. Non-hospitalized patients with lower IES-scores reported least problems on the HUI3 health domains. Symptoms indicative of PTSD (IES≥35) were associated with a mean utility loss of 0.33 in non-hospitalized patients (t=80.8; t<0.001) and 0.24 in hospitalized patients (t=15.9; t=0.001) (see figure 8.2).

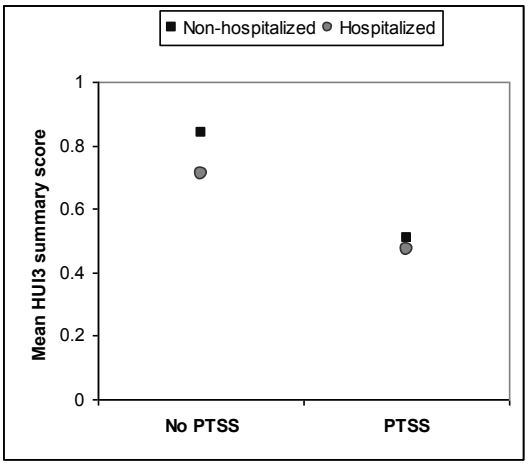

Figure 8.2 Mean HUI3 utility scores of hospitalized and non-hospitalized patients with and without symptoms indicative of post-traumatic stress disorder (PTSS) at 24-month

The models tested to predict HRQoL measured with EQ-5D and HUI3 were both statistically significant (EQ-5D: $F = 80.27$, $p < 0.001$; HUI3: $F = 118.55$, $p<0.001$). Table 8.4 shows that posttraumatic stress symptoms indicative of PTSD (IES≥35) are associated with decreased HRQoL, even after controlling for possible confounders.

Table 8.4 Predictors of health-related quality of life at 2-year follow-up[§]

	Predictors	R2	Standardized B	p
EQ-5D	Age		-0.023	0.335
	Sex		-0.134	<0.001
	Hospitalization		-0.246	<0.001
	Comorbidity		-0.238	<0.001
	PTSD (IES-score≥35)		-0.234	<0.001
		0.217		
HUI3	Age		-0.159	<0.001
	Sex		-0.058	0.008
	Hospitalization		-0.098	<0.001
	Comorbidity		-0.371	<0.001
	PTSD (IES-score≥35)		-0.211	<0.001
		0.278		

[§] Analysis based on stepwise multivariate regression analysis with demographics (age, sex) as block 1; comorbidity, hospitalization and severity level of the injury as step 2, and posttraumatic stress symptoms indicative of PTSD(IES-score≥35) as step 3.

Discussion

Two years after trauma, 4.1% of our comprehensive population of injury patients treated at an ED had an IES-score of 35 or higher, indicating evident symptoms of PTSD (IES≥35) [25]. Female gender, occupational injury and hospitalization were the strongest independent predictors of posttraumatic stress symptoms indicative of PTSD, which were associated with more problems on almost all domains of functional outcome and a considerable decrease of HRQoL in both non-hospitalized and hospitalized injury patients two years post-injury.

This study was not restricted to particular injury subgroups, whereas most studies regarding PTSD post-injury concern specific groups of injury patients such as victims of motor vehicle accidents or victims with severe injuries [4, 8, 11, 12, 19]. The high variety in injuries included in this study and the relatively large sample size allowed examination of the association of a number of injury characteristics and posttraumatic stress symptoms indicative of PTSD. It is difficult to compare results found in this study to PTSD prevalence rates reported in previous studies because of differences in studied injury population. Only one previous study conducted a long-term follow-up among a comprehensive sample of injury patients treated at an ED. The reported PTSD prevalence rate of this one-year follow-up study [16], was higher (10%) compared to the prevalence rate of PTSD that we found. This difference in prevalence rate might be explained by differences in patient populations and methods used to assess symptoms of PTSD. The follow-up time was with one-year much shorter than the two-year follow-up of the current study and PTSD symptoms were assessed using the Clinician-Administered PTSD Scale for DSM-IV rather than the IES. Furthermore, the respondents that were included were admitted to the trauma service for at least 24 hours. As a result, the injuries of the respondents included in the study of O'Donnell et al.[16] were on average much more severe compared to the injuries of the patients of the current study, which included all injury patients aged 15 years and older that attended the ED.

Other studies addressed PTSD prevalence rates among specific groups of injury victims. We can, however, compare subgroups of the injury patients included in the current study, such as victims of traffic accidents. Mayou et al. assessed PTSD among victims of traffic accidents treated at the emergency department and found a PTSD prevalence rate of 9% at one year post trauma and 12% at three years post trauma [5]. The PTSD prevalence rate of 5% among victims of traffic accidents found in the current study is slightly lower compared to this range of observed prevalence rates.

The results of our study indicate a strong association between female gender and PTSD indications. This association is in line with findings in the general literature [37] and has been reported by other studies on PTSD following injuries [3, 38-40]. Other independent predictors of PTSD indications at two-year follow-up were hospitalization, occupational injury, comorbid disease, and severity level of the sustained injuries.

Furthermore, we found that injury patients with posttraumatic stress symptoms indicative of PTSD reported significantly more problems on all EQ-5D and almost all HUI3 health

domains. A study that investigated HRQoL with EQ-5D among patients with PTSD following cardiac arrest reported similar findings [41]. Among adolescent victims PTSD was associated with impairments in Role/Social Behavioral, Role/Social Physical, Bodily Pain, General Behavior, Mental Health, and General Health Perceptions subscales of the 87-item Child Health Questionnaire [21]. The resulting EQ-5D and HUI3 utility scores of injury patients with PTSD found in the current study are approximately in the range of the utility scores that Holbrook et al. derived with the multi-attribute utility instrument Quality of Well-being scale (QWB) (0.58 – 0.62) [19]. Although the HUI3 instrument yielded significantly lower health utility scores compared to the EQ-5D, which accords with results of other studies [42-44], both HUI3 and EQ-5D showed that PTSD was associated with a mean utility loss of 0.17 – 0.25. This concurs with the utility loss of anxiety disorders social phobia, generalized anxiety disorder and agoraphobia [45].

It should be noted that Holbrook et al. focused on injury patients admitted to a trauma centre with a length of stay of more than 24 hours and patients injured due to unintentional and intentional injury, whereas the current study included all admitted injury patients to general and university hospitals who were injured due to unintentional injury. Moreover, Holbrook et al. used an IES-score greater than 24 to identify patients with PTSD, whereas in the current study a cut off of 35 was used. Evidence from studies on this matter suggests that to avoid overestimation of the number of cases with PTSD, an IES-score of greater than 35 is more appropriate [25, 46]. Using the DSM-IV as the diagnostic criteria for PTSD, a cut-off score of 35 produced sensitivity of .89, specificity of .94 [25]. With a cut-off point of 24, the sensitivity is 0.91 and the specificity 0.46 [47]. To avoid over diagnosing of PTSD in a comprehensive population with a relative low PTSD prevalence, it is important to use a high IES cutoff score that incurs a high specificity.

A limitation of this study was the low response rate of the follow-up questionnaires [22]. The 24-month follow-up questionnaire, which included the IES, was send only to those patients who responded to the preceding three follow-up questionnaires send at 2½, 5 and 9 months. This meant that only 21% of the patients of the initial sample selected for the follow-up study filled in the 24-month follow-up questionnaire. However, the data were adjusted for non-response and possible response bias, because the PTSD prevalence rates were calculated using data that were weighted with respect to the original sample size and composition by inverse probability weighting. For some aspects, such as the severity of sustained injuries, the adjustments of non-response could be improved, since injury severity scores were not available.

Evidence suggested that patients with very severe health problems are less likely to respond to a survey [48]. Differential underreporting by level of severity cannot be excluded, since we found a larger proportion of hospitalized patients among those with PTSD at 2 years post-injury. This could partly be caused by missing a larger share of the more severely injured hospitalized patients among those without PTSD (e.g. comatose patients). This may have led to a slight overestimation of the utility losses due to PTSD. However, severely injured patients are only a minor part of the total sample and PTSD

remained significantly associated to adverse HRQoL, even after adjustment for confounders including hospitalization status.
In the current study PTSD is measured at 24 months follow-up only. A longitudinal study on PTSD and HRQoL among injury patients might elucidate any causal relationship between PTSD and subsequent reduced HRQoL. Furthermore, the influence of earlier HRQoL on PTSD remains to be investigated.

We conclude that one in 25 patients admitted to an ED due to injuries of all causes and severity levels have an IES-score of 35 or higher, indicating evident symptoms of PTSD . Female gender, occupational injury and hospitalization were the strongest independent predictors of this disorder. Respondents with PTSD indications report significantly more problems on almost all HRQoL domains. Additionally, the results of the current study showed that posttraumatic stress symptoms indicative of PTSD are associated with decreased HRQoL even after correction for possible confounders such as comorbidity. PTSD seems a major barrier for full recovery of injury patients of even minor levels of severity, and the development and evaluation of ED based policies for its early diagnosis and treatment should therefore be stimulated.

References

1. Blanchard, E.B., et al., Psychophysiology of posttraumatic stress disorder related to motor vehicle accidents: replication and extension. J Consult Clin Psychol, 1996. 64(4): p. 742-51.
2. Mayou, R., S. Tyndel, and B. Bryant, Long-term outcome of motor vehicle accident injury. Psychosom Med, 1997. 59(6): p. 578-84.
3. Holbrook, T.L., et al., Gender differences in long-term posttraumatic stress disorder outcomes after major trauma: women are at higher risk of adverse outcomes than men. J Trauma, 2002. 53(5): p. 882-8.
4. Ehlers, A., R.A. Mayou, and B. Bryant, Psychological predictors of chronic posttraumatic stress disorder after motor vehicle accidents. J Abnorm Psychol, 1998. 107(3): p. 508-19.
5. Mayou, R.A., A. Ehlers, and B. Bryant, Posttraumatic stress disorder after motor vehicle accidents: 3-year follow-up of a prospective longitudinal study. Behav Res Ther, 2002. 40(6): p. 665-75.
6. Meiser-Stedman, R., et al., Maladaptive cognitive appraisals mediate the evolution of posttraumatic stress reactions: A 6-month follow-up of child and adolescent assault and motor vehicle accident survivors. J Abnorm Psychol, 2009. 118(4): p. 778-87.
7. Yasan, A., et al., Predictive factors for acute stress disorder and posttraumatic stress disorder after motor vehicle accidents. Psychopathology, 2009. 42(4): p. 236-41.
8. McKibben, J.B., et al., Acute Stress Disorder and Posttraumatic Stress Disorder: A Prospective Study of Prevalence, Course, and Predictors in a Sample With Major Burn Injuries. J Burn Care Res, 2008. 29(1): p. 22-35.
9. Corry, N.H., B. Klick, and J.A. Fauerbach, Posttraumatic stress disorder and pain impact functioning and disability after major burn injury. J Burn Care Res, 2010. 31(1): p. 13-25.
10. Dyster-Aas, J., et al., Major depression and posttraumatic stress disorder symptoms following severe burn injury in relation to lifetime psychiatric morbidity. J Trauma, 2008. 64(5): p. 1349-56.

11. Hepp, U., et al., Coping with serious accidental injury: a one-year follow-up study. Psychother Psychosom, 2005. 74(6): p. 379-86.
12. Schnyder, U., et al., Incidence and prediction of posttraumatic stress disorder symptoms in severely injured accident victims. Am J Psychiatry, 2001. 158(4): p. 594-9.
13. Quale, A.J., et al., Severity of injury does not have any impact on posttraumatic stress symptoms in severely injured patients. Injury, 2009. 40(5): p. 498-505.
14. Myhren, H., et al., Posttraumatic stress, anxiety and depression symptoms in patients during the first year post intensive care unit discharge. Crit Care, 2010. 14(1): p. R14.
15. O'Donnell M, L., et al., Posttraumatic Stress Disorder After Injury: Does Admission to Intensive Care Unit Increase Risk? J Trauma, 2010.
16. O'Donnell, M.L., et al., Psychiatric morbidity following injury. Am J Psychiatry, 2004. 161(3): p. 507-14.
17. Zatzick, D., Posttraumatic stress, functional impairment, and service utilization after injury: a public health approach. Semin Clin Neuropsychiatry, 2003. 8(3): p. 149-57.
18. Zatzick, D.F., et al., Posttraumatic stress disorder and functioning and quality of life outcomes in a nationally representative sample of male Vietnam veterans. Am J Psychiatry, 1997. 154(12): p. 1690-5.
19. Holbrook, T.L., et al., Perceived threat to life predicts posttraumatic stress disorder after major trauma: risk factors and functional outcome. J Trauma, 2001. 51(2): p. 287-92; discussion 292-3.
20. Landolt, M.A., et al., Post-traumatic stress impacts on quality of life in children after road traffic accidents: prospective study. Aust N Z J Psychiatry, 2009. 43(8): p. 746-53.
21. Zatzick, D.F., et al., Association between posttraumatic stress and depressive symptoms and functional outcomes in adolescents followed up longitudinally after injury hospitalization. Arch Pediatr Adolesc Med, 2008. 162(7): p. 642-8.
22. Polinder, S., et al., Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma, 2007. 62(1): p. 133-41.
23. Curtis, L.H., et al., Using inverse probability-weighted estimators in comparative effectiveness analyses with observational databases. Med Care, 2007. 45(10 Supl 2): p. S103-7.
24. Horowitz, M., N. Wilner, and W. Alvarez, Impact of Event Scale: a measure of subjective stress. Psychosom Med, 1979. 41(3): p. 209-18.
25. Wohlfarth, T.D., et al., Screening for Posttraumatic Stress Disorder: an evaluation of two self-report scales among crime victims. Psychol Assess, 2003. 15(1): p. 101-9.
26. van der Ploeg, E., et al., Construct validation of the Dutch version of the impact of event scale. Psychol Assess, 2004. 16(1): p. 16-26.
27. Drummond, M., et al., Methods for the Economic Evaluation of Health Care Programmes. 1997, Oxford, U.K.: Oxford Medical Publications.
28. Rabin, R. and F. de Charro, EQ-5D: a measure of health status from the EuroQol Group. Ann Med, 2001. 33(5): p. 337-43.
29. Coons, S.J., et al., A comparative review of generic quality-of-life instruments. Pharmacoeconomics, 2000. 17(1): p. 13-35.
30. Van Beeck, E.F., et al., Guidelines for the conduction of follow-up studies measuring injury-related disability. J Trauma, 2007. 62(2): p. 534-50.
31. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group Health Policy, 1990. 16(3): p. 199-208.

32. Feeny, D., et al., Multi-attribute health status classification systems. Health Utilities Index. Pharmacoeconomics, 1995. 7(6): p. 490-502.
33. CBS, Permanent onderzoek leefsituatie. . 2003, Statistics Netherlands (CBS): Voorburg/Heerlen.
34. Mosby, Mosby's Medical Dictionary,. 8th ed. 2009: Elsevier Health Sciences.
35. Rubin, D.B. and N. Schenker, Multiple imputation in health-care databases: an overview and some applications. Stat Med, 1991. 10(4): p. 585-98.
36. Polinder, S., et al., International variation in clinical injury incidence: exploring the performance of indicators based on health care, anatomical and outcome criteria. Accid Anal Prev, 2008. 40(1): p. 182-91.
37. Olff, M., et al., Gender differences in posttraumatic stress disorder. Psychol Bull, 2007. 133(2): p. 183-204.
38. Breslau, N., et al., Sex differences in posttraumatic stress disorder. Arch Gen Psychiatry, 1997. 54(11): p. 1044-8.
39. Mayou, R. and B. Bryant, Consequences of road traffic accidents for different types of road user. Injury, 2003. 34(3): p. 197-202.
40. Fullerton, C.S., et al., Gender differences in posttraumatic stress disorder after motor vehicle accidents. Am J Psychiatry, 2001. 158(9): p. 1486-91.
41. Gamper, G., et al., Life after death: posttraumatic stress disorder in survivors of cardiac arrest--prevalence, associated factors, and the influence of sedation and analgesia. Crit Care Med, 2004. 32(2): p. 378-83.
42. Grutters, J.P., et al., Choosing between measures: comparison of EQ-5D, HUI2 and HUI3 in persons with hearing complaints. Qual Life Res, 2007. 16(8): p. 1439-49.
43. Marra, C.A., J.M. Esdaile, and A.H. Anis, Practical pharmacogenetics: the cost effectiveness of screening for thiopurine s-methyltransferase polymorphisms in patients with rheumatological conditions treated with azathioprine. J Rheumatol, 2002. 29(12): p. 2507-12.
44. Oostenbrink, R., A.M. HA, and M.L. Essink-Bot, The EQ-5D and the Health Utilities Index for permanent sequelae after meningitis: a head-to-head comparison. J Clin Epidemiol, 2002. 55(8): p. 791-9.
45. Saarni, S.I., et al., Impact of psychiatric disorders on health-related quality of life: general population survey. Br J Psychiatry, 2007. 190: p. 326-32.
46. Neal, L.A., et al., Convergent validity of measures of post-traumatic stress disorder in a mixed military and civilian population. J Trauma Stress, 1994. 7(3): p. 447-55.
47. Shalev, A.Y., et al., Prospective study of posttraumatic stress disorder and depression following trauma. Am J Psychiatry, 1998. 155(5): p. 630-7.
48. Hebert, R., et al., Refusal and information bias associated with postal questionnaires and face-to-face interviews in very elderly subjects. J Clin Epidemiol, 1996. 49(3): p. 373-81.

9

Beyond the neglect of psychological consequences: posttraumatic stress disorder increases the non-fatal burden of injury by more than 50%

Juanita A. Haagsma, Suzanne Polinder, Hidde Toet, Martien Panneman, Arie H. Havelaar, Gouke J. Bonsel, Ed F. van Beeck

Accepted by Injury Prevention.

Abstract

Introduction Psychological consequences such as posttraumatic stress disorder (PTSD) are currently neglected in burden of injury calculations. The aim of this study was to assess the disease burden of PTSD due to unintentional injury and compare this with health loss to physical injury consequences.

Methods From literature sources, the prevalence rate of PTSD at four follow-up periods (<3 months, 3-6 months, 7-12 months and >12 months) was estimated. The uncertainty of the estimated PTSD prevalence rate was modelled by a Baysian approach. The prevalence rates were then linked to national data on unintentional injury, disability weights and duration to estimate the incidence and Disability Adjusted Life Years (DALY) resulting from PTSD in addition to physical injury consequences.

Results The data suggest that PTSD prevalence rate among injury victims decreases over time. The average PTSD prevalence rate at <3 months follow-up was 21% (90% credibility interval (CI) 17-24%) for patients presenting at the ED and 30% (90% CI 27-33%) for patients that were hospitalized, after 12 months tapering down to 4% (90% CI 3-5%) and 6% (90% CI 4-10%) respectively. These estimates translate into 191,000 (90% CI 161,000-222,000) cases of PTSD per year in the Dutch population (1.2%) due to unintentional injury. Including PTSD increases the non-fatal burden of disease of unintentional injuries by 53% (from 116,000 to 178,000 DALYs (90% CI 150,000-217,000).

Conclusions Ignoring PTSD in burden of injury studies results in a considerable underestimation of the burden of injury. This may impact resource allocation and the identification of important prevention priorities.

Introduction

Following the 1996 Global Burden of Disease study, injury researchers have used the Disability Adjusted Life Year (DALY) in burden of disease and injury studies to facilitate effective and efficient policy-making on injury prevention and surveillance [1]. The DALY is a metric that integrates the effect of mortality and morbidity and disability so that they can be considered at the same time. This simplifies comparison of health outcomes between subgroups of a population. A prerequisite for comparison is that the burden of disease or injury comprises a complete assessment of the health effects due to a certain cause. In previous burden of injury studies, long-term injury consequences such as permanent disability have been included; however, with important restrictions as was pointed out by Lyons et al. [2]. In their paper, Lyons et al. presented a framework that comprises the injury burden across individual, family and societal domains, including psychological consequences. So far, only physical consequences were accounted for, whereas psychological consequences such as posttraumatic stress disorder (PTSD), depression and other maladaptive syndromes – undisputed consequences of high impact once they occur – have not previously been considered. Many studies have revealed high prevalence of disabling psychological consequences among injury victims and this neglect underestimates the injury problem by health policy makers, affecting resource allocation and the identification of important prevention opportunities [3-8].

A common psychological problem causally related to the injury event is PTSD. PTSD may result from any event that involves an injury, or threatened or actual death (of others). PTSD meets the definition of an injury sequela, i.e. a pathological condition resulting from a prior injury, disease or attack. Along with its high prevalence among injury victims of all severity levels [9-12], its large effect on health-related quality of life, and the usually long term presence of psychiatric consequences, PTSD following injury might contribute a substantial burden that is currently neglected in estimates of the burden of injury.

In this study, we aimed to

1) Assess the prevalence of PTSD among injury patients of all unintentional causes,

2) Estimate the disease burden in DALYs resulting from PTSD due to unintentional injury in The Netherlands and compare the resulting health loss to physical injury consequences.

Methods

Study design

A literature review was conducted to obtain data on the prevalence rate of PTSD among injury patients. In order to include patients with minor and major severity levels, we focused on follow-up studies of samples treated at an emergency department (ED) and/or admitted to hospital. Follow-up period is defined as the time between treatment at

the ED or admission to the hospital and measurement of PTSD symptoms. To estimate the prevalence rate of PTSD at various follow-up periods, the reported data of the studies were subdivided into three groups; <3 months, 3-6 months, 7-12 months and >12 months. The PTSD prevalence rates were calculated for each of these follow-up periods. The uncertainty of the estimated PTSD prevalence rates at these four follow-up periods was modelled by using a Baysian approach [13]. The PTSD prevalence rate was considered as a binomial fraction in which the uncertainty was modelled as a Beta (s,f) distribution. Here, s is the number of persons who developed PTSD and f is the number of persons who did not. Using this approach, the 90% credibility interval (CI) is similar to that estimated by conventional means, with the advantage that a full uncertainty distribution can be simulated.
As a next step, the number of patients that visited the ED due to unintentional injury in the Netherlands was obtained from registry systems. The estimated prevalence rate with uncertainty by follow-up moment was multiplied by the number of injury cases to estimate the incidence of cases with PTSD in the Netherlands. The burden of PTSD following injury was then estimated by combining these incidence estimates with appropriate disability weights and estimates for the duration of PTSD.

Literature review

To obtain data on the prevalence rate of PTSD among injury patients, a literature review was conducted of published studies to date that explored PTSD among injury patients. This literature review comprised a search of the database MEDLINE and Psychinfo that was restricted to articles published in English from and including 1990 through March 2010, using the following search terms: posttraumatic stress disorder, injury, follow-up, psychiatric, psychological, incidence and prevalence. Excluded were review articles, articles on combat-related injury, injury caused by (natural) disaster or war, criminal violence, partner violence and sexual assault. The resulting articles were screened to identify articles that met the predefined selection criteria. Reference lists of the included articles were also examined to identify cited articles not captured by MEDLINE and Psychinfo. Finally, full text articles that met the selection criteria were scrutinized to assess methodological quality. We only included studies that addressed:

- adult injury patients of all ages, both male and female
- injury patients treated at the ED and/or admitted to hospital
- injuries of all causes (traffic, occupation, home and leisure, sports and violence)
- all types of injuries.

These studies were supplemented by further analysis of unpublished data from a published Dutch follow-up study which measured PTSD among 1,781 injury patients 24 months after visiting an ED department of a hospital [14].

Burden of disease calculation

The incidence of injury patients aged 15 years or older treated at the ED was derived from the Dutch Injury Surveillance System, a continuous registry of intentional and unintentional injuries of 17 hospitals in the Netherlands [14]. Registered patients are victims of intentional and unintentional injury. This study was deliberately restricted to victims of unintentional injury (home and leisure, traffic, occupational and sport). The sustained injuries varied from minor to severe injury, single and multiple injury and hospitalized and non-hospitalized patients. The estimated incidence of PTSD was calculated by combining the incidence of injury and the calculated mean prevalence rate of PTSD.

We calculated the mean prevalence rate of PTSD for patients treated at the ED and hospitalized patients separately. Subsequently, these prevalence rates were used to calculate the burden of disease due to PTSD, expressed in DALY, in these two groups.

The DALY aggregates mortality, expressed in years of life lost (YLL), and morbidity, expressed in years lived with disability (YLD). Regarding PTSD, the mortality component of the DALY is not taken into account as it is not associated with increased mortality. To calculate YLD resulting from PTSD, the incidence of PTSD was combined with a disability weight and duration of the condition. The disability weight is a value anchored between 0 and 1 that is assigned to living with the condition. This value reflects the impact of a specific health condition on health-related quality of life and it is commonly based on the preferences of a panel of judges. We adopted the PTSD disability weight derived from an empirical disability weights study [15], a Dutch study performed in 2006 that derived disability weights for 44 health states from a panel of lay men (n=107) recruited from the general population. The 44 health states included health states related to infectious disease, work-related health states and psychiatric health states. The procedure that was followed to derive disability weights for these health states was similar to the study performed in 2004 that derived disability weights for 45 permanent and temporary consequences of injury [16]. The health state descriptions, or vignettes, included the disease label, disease specific information concerning the symptoms, functional health status in generic terms and information regarding the duration of the symptoms. The panel of laymen valued these vignettes using the time trade-off valuation technique. In this study a renewed methodology was applied that focused especially on obtaining and improving disability weights for functional losses of a temporary and complex nature. The procedures used to derive the panel study disability weights of the 2004 (injury) and 2006 (non-injury) disability weights studies were described in detail in previously published papers [15, 16]. The population panel valued several PTSD health states, with a duration varying from 3 months, 6 months and one whole year. A group of patients may have symptoms for more than one year and to calculate the YLD due to PTSD in this group of patients, information is needed on the maximum duration of the symptoms. The few longitudinal studies that have investigated the course of PTSD among injury patients, reported mean durations of PTSD symptoms of 6 through 8 years

[17, 18]. Regarding this study, we assumed that the group of patients with long-term PTSD symptoms (>1 years) will experience PTSD symptoms for on average 5.5 years, with a minimum duration of 3 years and maximum duration of 8 years.
The existing YLD for short and long-term physical consequences of injury were calculated using incidence data from the Dutch Injury Surveillance System and the set of 45 annual profile disability weights for permanent and temporary consequences of injury [16]. This approach was used in the National Public Health Forecast of the Netherlands [19],which so far has neglected the impact of psychological consequences of injury on population health. The neglect was due to the fact that at the release of this Dutch policy document estimates on the prevalence of PTSD following injury and PTSD disability weights tailored to the prevalence data were not yet available.

Results

Literature review

We found 125 studies that investigated the incidence of PTSD among injury patients. Regarding these studies, 101 were restricted to specific subgroups of injury patients, namely patients that sustained injury due to specific injury cause (e.g. traffic injury or violence), patients with specific injuries (e.g. burns or brain injury), specific groups of injury patients (e.g. children, elderly or male patients) or patients with very severe injuries (e.g. admitted to the ICU). One study did not report PTSD prevalence rates. Seven studies reported on PTSD in the same sample of injury patients of other studies included in our literature review.
Sixteen studies investigated adult injury patients presenting at the ED and/or hospitalized due to unintentional injury of all unintentional causes. In addition, unpublished data could be analysed from a published Dutch follow-up study which measured PTSD 24 months after visiting an ED department (resulting in either hospitalisation or discharge to the home environment). This brings the total number of studies to 17. Regarding these 17 studies, follow-up periods ranged from one week through 24 months and PTSD prevalence rates ranged from 2% through 50%. Table 9.1 presents an overview of the findings of these 17 follow up studies.

PTSD prevalence rates

ED treatment - The average PTSD prevalence rate at <3 months follow-up was PTSD prevalence rate at <3 months follow-up was 21% with a 90% CI of 17-24%. At 3-6 months, 7-12 months and >12 months the PTSD prevalence rates regarding patients treated at the ED were 12% (CI 10-14%), 9% (CI 8-12%) and 4% (CI 3-5%), respectively.

Table 9.1 Study characteristics, methodological aspects and PTSD rates of the 17 included studies in the review on injury patients treated at the Emergency Department and/or admitted to hospital (aged >15 y, all causes)

Study and country	Study population	Assessment tool*	PTSD rates
Bryant et al. [20]	n=1084 (1477), age 16-70	CAPS	3 months: 10%
Australia	Hospitalized		12 months: 10%
Holbrook et al. [21]	n=824 (1048), age 18 or older	SCID	6 months: 32%
USA	Hospitalized		
Mason et al. [22]	n=653	IES-R	3 months: 37%
UK	Hospitalized		
Matthews et al. [23]	n=69 (219), age 18-60	PCL	8 months: 19%
Australia	Hospitalized		
Mellman et al. [24]	n=50 (83)	CAPS, SCID	6 weeks: 24%
USA	Hospitalized		
Michaels et al [25]	n=247, age 18 or older	MISS	6 months: 41%
USA	Hospitalized		12 months: 38%
Micheals et al. [26]	n= 56 (126)	MISS, IES	5 months: 38%
USA	Hospitalized		
O'Donnell et al. [6]	n=337 (419), age 18-70	CAPS	3 months: 9%
Australia	ED and hospitalized		12 months: 10%
Richmond et al. [27]	n=123 (228), age 18-80	IES	1 week: 32%
USA	Hospitalized		12 weeks: 50%
Schnyder et al. [28]	n=106 (135), age 18-70	IES, CAPS	2 weeks: 5%
Switzerland	Hospitalized		1 year: 2%
Schnyder et al [29]	n=255 (386), age 18-65	CAPS	6 months: 3%
Switserland	Hospitalized		
Shalev et al. [11]	n=211 (420), aged 16-65	IES, CAPS,	1 month: 30%
Israel	ED and hospitalized	MISS	4 months: 18%
Shalev et al. [30]	n=51 (72), aged 18-60	IES, SCID,	6 months: 26%
Israel	Hospitalized	MISS	
Turpin et al. [12]	n=142 (2,818), aged 16-65	PDS	<2 weeks: 14%
UK	ED and hospitalized		10-12 weeks: 7%
			24-26 weeks: 10%
Zatzick et al. [8]	n=101 (156), age 14-65	PCL	12 months: 30%
USA	Hospitalized		
Zatzick et al. [31]	n=2707 (5043), age 18-84	PCL	12 months: 21%
USA	Hospitalized		
Polinder et al. [14]	n=1,781 (8,564), aged >15	IES	24 months: 4%
The Netherlands	ED and hospitalized		

* CAPS = Clinician-Administered PTSD Scale, SCID = Structured Clinical Interview for DSM-IV, IES = Impact of Event Scale, IES-R = revised impact of event scale, MISS = Mississippi Scale for Combat-Related PTSD, PCL = PTSD Checklist, PDS = Posttraumatic Stress Diagnostic Scale

Comparison of the PTSD prevalence rate with the PTSD prevalence rate of the subsequent follow-up period indicated that this decrease was significant (comparison of <3months and 3-6 months PTSD prevalence rate: z=3.9, p<0.001; 3-6 months and 7-12 months: z=1.9, p=0.23; 7-12 months and >12 months z=5.0, p<0.001).
Hospitalization - The average PTSD prevalence rate regarding patients admitted to hospital at <3 months, 3-6 months, 7-12 months and >12 months follow-up were 30% (CI 27-33%), 24% (CI 22-25%), 19% (CI 18-20%) and 6% (CI 4-10%), respectively. The data suggested that this decrease was significant (comparison of <3months and 3-6 months PTSD prevalence rate: z=2.4, p<0.001; 3-6 months and 7-12 months: z=3.8, p=0<0.001; 7-12 months and >12 months z=15.5, p<0.001).
The PTSD prevalence rates of hospitalized patients were significantly higher in the first year post-injury compared to non-hospitalized patients (comparison <3months: z=3.1, p<0.001; 3-6 months: z=7.0, p<0.001; 7-12 months: z=5.6, p<0.005; >12 months z=1.2, p=0.24). The calculated PTSD prevalence with uncertainty limits per follow-up period are shown in figure 9.1.

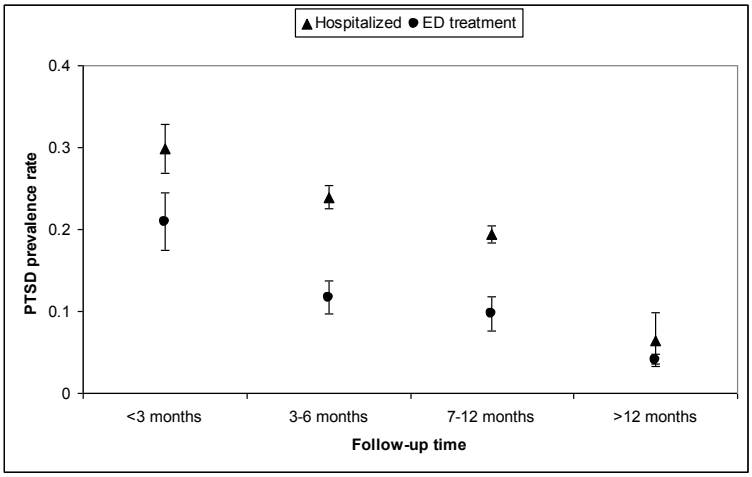

Figure 9.1 Prevalence rate of posttraumatic stress disorder (PTSD) following unintentional injury treated at the Emergency Department (ED) and admitted to hospital. Error bars express the 90% credibility interval of the estimated prevalence rate.

Burden of disease calculation

The data sources showed that in 2007 in the Netherlands (population size 16.4 million inhabitants) 870,000 patients visited the ED due to unintentional injury (i.e. 5% of the Dutch population) [14]. Of these patients, 105,000 were admitted to hospital. An estimated 191,000 (CI 161,000-222,000) of these injury patients attending the ED and/or admitted to hospital would have developed PTSD. The estimated number of YLDs lost due to PTSD following unintentional injury is shown in table 9.2. The disease burden of

PTSD following unintentional injury is in total an estimated 61,936 YLDs (CI 34,369-100,759), increasing the non-fatal burden of unintentional injury in the Netherlands with 53% from 116,092 to 178,028 YLDs (CI 150,461-216,851; see figure 9.2).

Table 9.2 Incidence, duration and disease burden of posttraumatic stress disorder (PTSD) following unintentional injury treated at the ED and or admitted to hospital (most likely estimates)

PTSD health state	Prevalence rate	n	Disability weight*	Duration	YLD	YLD per 100,000
ED treatment						
PTSD, 3 months	9.1%	70,000	0.05	*	3,499	
PTSD, 6 months	2.1%	16,000	0.133	*	2,107	
PTSD, whole year	5.6%	43,000	0.211	*	9,098	
PTSD, > 24 months	4.0%	31,000	0.211	5.5	35,585	
ED total	*20.8%*	*160,000*			*50,288*	*307.4*
Hospitalized						
PTSD, 3 months	5.9%	6,200	0.05	*	309	
PTSD, 6 months	4.5%	4,700	0.133	*	628	
PTSD, whole year	13.0%	14,000	0.211	*	2,870	
PTSD, > 24 months	6.4%	6,800	0.211	5.5	7,841	
Hospitalized total	*29.8%*	*31,000*			*11,648*	*71.2*
Total	26.6%	191,000			61,936	378.6

* annual profile disability weights

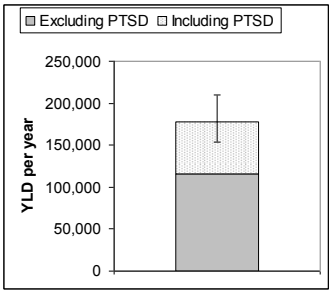

Figure 9.2 Total disease burden (in YLDs) due to physical consequences and PTSD following unintentional injury treated at the ED and/or admitted to hospital. Error bars express the uncertainty of the estimated YLD

Discussion

From literature sources, we estimated that the average PTSD prevalence rate at <3 months follow-up was 21% (CI 17-24%) for patients presenting at the ED and 30% (CI 27-33%) for patients that were hospitalized, after 12 months tapering down to 4% (CI 3-5%) and 6% (CI 4-10%) respectively. These estimates translate into 191,000 (CI 161,000-222,000) cases of PTSD following unintentional injury (1.2% of the Dutch population). Including PTSD as injury burden in addition to physical consequences increases the non-fatal burden of disease of unintentional injuries by more than 50%.

Previous studies suggested that PTSD is common among injury victims, exerting a high impact on health-related quality of life [10, 21, 32-34]. Our study integrated these aspects using the DALY to gain insight in the disease burden associated with PTSD following unintentional injury. For this integration little information considering key aspects was available. These aspects concerned PTSD prevalence among injury patients treated at the ED and/or admitted to hospital, the course of PTSD and the maximum duration of PTSD symptoms. Regarding these aspects conservative assumptions have been made using international literature. The uncertainty incurred by these assumptions was made explicit by statistical simulation.

The limited information on these three key aspects (i.e. PTSD prevalence, course and maximum duration) might have affected our results. Firstly, the data that were used in the current study to estimate the occurrence of PTSD were retrieved from existing literature. Several follow-up studies were used for this purpose and many of these studies used self-report questionnaires, such as the IES, to identify cases with PTSD. Self-report questionnaire are not designed to diagnose mental disorders according to the DSM-IV (the fourth edition of the diagnostic and statistical manual for psychiatric disorders). As a result, cases that were identified as having posttraumatic stress symptoms indicative of PTSD might not meet the DSM-IV criteria of clinical PTSD, and conversely. If we restrict the literature review to studies that used structured clinical interviews to identify cases with PTSD, the PTSD prevalence rates at <3 months, 3-6 months and 7-12 months would be 22%, 18% and 10% respectively (hospitalized and non-hospitalized combined). These prevalence rates are lower compared to the prevalence rates that are based on studies that use self-report and structural clinical interview (<3 months 27%; 3-6 months 21%; 7-12 months 18%). This implies that by including studies that use self-report questionnaires to identify PTSD we may have overestimated the number of cases with PTSD and consequently the burden of disease of PTSD following unintentional injury.

Secondly, regarding the duration of PTSD among injury victims, a long-term follow-up study among patients with motor vehicle traffic injury showed that the PTSD prevalence rate remained more or less the same at 3 months, 1 year and 5 years post-injury; five years after exposure 10% of the subjects had persistent PTSD symptoms [35]. This suggests that our assumption of a maximum PTSD duration of eight years might be on the cautious side, underestimating the actual burden of PTSD.

We did not take into account that PTSD symptoms fluctuate over time and that in case of long-term PTSD the symptoms may not have existed since day one. To allow correction for such fluctuations, detailed epidemiological follow-up data of injury patients are needed. Regarding the current study, it was not possible to correct for these fluctuations in PTSD symptoms, since studies included in the review did not report details regarding the dynamics of PTSD symptoms in their study population.

Furthermore, in the current study PTSD prevalence rates derived from literature were applied to cases seen in the Dutch Injury Surveillance system to estimate the number of PTSD cases following injury in the Netherlands. However, there are possible differences between cases seen in the Dutch Injury Surveillance system and the smaller number of cases included in studies from the literature; differences such as severity of the injury. These differences in cases may affect the risk of developing PTSD post-injury. Therefore, applying the estimated PTSD prevalence rates derived from literature to cases of the Dutch Injury Surveillance may over- or underestimated the number of cases with PTSD. To avoid this, we calculated PTSD prevalence rates and disease burden for hospitalized and non-hospitalized patients separately.

When comparing the total disease burden due to injury (YLL and YLD) to that of diseases in the Netherlands, home and leisure injuries and traffic injuries rank at number 11 and 17, respectively. Cardiovascular disease tops the list with 339,700 DALYs and with 101,500 DALYs dementia ranks at number 10. If PTSD is included in the burden of injury calculations, home and leisure injury will rise in ranking to number 8, with 127,000 DALYs and traffic injury will rise to number 13, with 78,000 DALYs.

Ranking of injury among major causes of health loss would change even more if this study was not restricted to PTSD, since other psychological consequences such as depression and travel anxiety are also frequently occurring consequences of injury [6, 11, 24, 36]. These findings suggest that our burden of injury estimates are still conservative and emphasize the importance of including psychological consequences to improve burden of injury estimations as well as other outcomes associated with injuries [37].

However, PTSD and other psychiatric disorders that may be caused by injuries are highly prevalent in the general population [38, 39]. In the Netherlands, the estimated prevalence of current PTSD in the total population was 3.8% [38]. This means that the estimated cases of PTSD following unintentional injury are approximately a third of the total estimated prevalence of PTSD in the Netherlands. However, new evidence revealed that approximately half of the injury patients diagnosed with PTSD experienced symptoms before the injury [20]. To unravel the intricate combinations of factors that cause PTSD and to aid in the attribution of its disease burden, more insight in the relation between injury and psychological disorders is needed. If this aspect is not addressed, the disease burden of PTSD cases may be attributed to multiple causes and consequently double counted in burden of disease studies.

Regarding the burden of injury calculation, the method used to assess disability weights is also an issue of importance. Several burden of injury studies used disability weights that were assessed with generic health status measures such as the EQ-5D [37, 40, 41]. These

measures allow patients to describe their health state by choosing a functional level for each generic health domain. The reported health state is then converted into a disability weight which by definition fits within the 0-1 range. Regarding the EQ-5D, the health profile is rendered by five attributes, i.e. mobility, self care, usual activities, pain/discomfort and anxiety/depression. If patient reported data are used for the assessment of disability weights, the disability weights may already include (part of) the impact of PTSD. We decided to use the panel derived disability weights, which involves a laymen panel to value health state descriptions, or vignettes. Using these disability weights avoids the issue of double counting that arises when disability weights derived with generic health status measures are used.
Furthermore, it should be noted that in the current study it was assumed that PTSD prevalence rates are consistent across countries. However, evidence suggests that PTSD prevalence rates differ widely across countries [42]. These differences have also been found regarding PTSD among injury patients across countries. In countries such as Switzerland, Japan and Australia low PTSD prevalence rates have been found, whereas in the United States of America much higher rates have been reported, even though the same methodology was used [6, 8, 34, 43].

We conclude that PTSD is a frequent sequel of injury and that the burden of PTSD is considerable compared to other outcomes of injury. The PTSD prevalence rate among injury patients declines over time, however, in some patients it remains persistent. Ignoring PTSD in burden of injury estimates results in a large underestimation of the size of the burden of injury. To satisfy the aspirations of burden of disease studies, namely identifying priorities in medical resource allocation and targeting interventions, sequelae of injury, such as PTSD, should be included, given that the evidence on the association is conclusive. Our disease burden estimates are likely to be conservative and further emphasize the importance of including psychological consequences to improve burden of injury estimations.

Implications for Prevention

The burden of disease concept is used increasingly for priority setting in healthcare and prevention. However, calculations of the burden of injury have important restrictions, since they focus on physical outcomes only, whereas injury is associated with a range of outcomes, including psychopathological consequences such as PTSD. The current study showed that neglecting psychological consequences in burden of injury estimations leads to a major underestimation and this may affect resource allocation and the identification of important prevention priorities.

References

1. Murray, C.J.L., A.D. Lopez, and C.D.e. Mathers, Summary measures of population health: concepts, ethics, measurement and applications. 2002, Geneva: World Health Organization.
2. Lyons, R.A., et al., The Injury List Of All deficits (LOAD) Framework - conceptualising the full range of deficits and adverse outcomes following injury and violence. Int J Inj Contr Saf Promot, 2009: p. 1-15.
3. Ehlers, A., R.A. Mayou, and B. Bryant, Psychological predictors of chronic posttraumatic stress disorder after motor vehicle accidents. J Abnorm Psychol, 1998. 107(3): p. 508-19.
4. Ellis, A., G. Stores, and R. Mayou, Psychological consequences of road traffic accidents in children. Eur Child Adolesc Psychiatry, 1998. 7(2): p. 61-8.
5. Mayou, R.A., A. Ehlers, and B. Bryant, Posttraumatic stress disorder after motor vehicle accidents: 3-year follow-up of a prospective longitudinal study. Behav Res Ther, 2002. 40(6): p. 665-75.
6. O'Donnell, M.L., et al., Psychiatric morbidity following injury. Am J Psychiatry, 2004. 161(3): p. 507-14.
7. van Reekum, R., et al., Psychiatric disorders after traumatic brain injury. Brain Inj, 1996. 10(5): p. 319-27.
8. Zatzick, D., Posttraumatic stress, functional impairment, and service utilization after injury: a public health approach. Semin Clin Neuropsychiatry, 2003. 8(3): p. 149-57.
9. Blanchard, E.B., et al., Psychophysiology of posttraumatic stress disorder related to motor vehicle accidents: replication and extension. J Consult Clin Psychol, 1996. 64(4): p. 742-51.
10. Holbrook, T.L., et al., Long-term posttraumatic stress disorder persists after major trauma in adolescents: new data on risk factors and functional outcome. J Trauma, 2005. 58(4): p. 764-9; discussion 769-71.
11. Shalev, A.Y., et al., Prospective study of posttraumatic stress disorder and depression following trauma. Am J Psychiatry, 1998. 155(5): p. 630-7.
12. Turpin, G., M. Downs, and S. Mason, Effectiveness of providing self-help information following acute traumatic injury: randomised controlled trial. Br J Psychiatry, 2005. 187: p. 76-82.
13. Vose, D., Risk Analysis, A Quantative Guide 2000, Chichester: John Wiley and Sons.
14. Polinder, S., et al., Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma, 2007. 62(1): p. 133-41.
15. Haagsma, J.A., et al., Disability Adjusted Life Years and minimal disease: application of a preference-based relevance criterion to rank enteric pathogens. Popul Health Metr, 2008. 6: p. 7.
16. Haagsma, J.A., et al., Novel empirical disability weights to assess the burden of non-fatal injury. Inj Prev, 2008. 14(1): p. 5-10.
17. Jeon, H.J., et al., Partial versus full PTSD in the Korean community: prevalence, duration, correlates, comorbidity, and dysfunctions. Depress Anxiety, 2007. 24(8): p. 577-85.
18. Taylor, S., Advances in the treatment of posttraumatic stress disorder. Cognitive-behaveriol perspective. 2004, New York: Springer.
19. de Hollander, A.E.M., et al., Care for health: The 2006 Dutch Public Health Status and Forecasts Report. 2007, National Institute for Public Health and the Environment: Bilthoven.
20. Bryant, R.A., et al., The psychiatric sequelae of traumatic injury. Am J Psychiatry, 2010. 167(3): p. 312-20.
21. Holbrook, T.L., et al., Perceived threat to life predicts posttraumatic stress disorder after major trauma: risk factors and functional outcome. J Trauma, 2001. 51(2): p. 287-92; discussion 292-3.
22. Mason, S., et al., The development of a clinically useful tool for predicting the development of psychological disorder following injury. Br J Clin Psychol, 2009. 48(Pt 1): p. 31-45.
23. Matthews, L.R., L.M. Harris, and S. Cumming, Trauma-related appraisals and coping styles of injured adults with and without symptoms of PTSD and their relationship to work potential. Disabil Rehabil, 2009. 31(19): p. 1577-83.

24. Mellman, T.A., et al., Predictors of post-traumatic stress disorder following severe injury. Depress Anxiety, 2001. 14(4): p. 226-31.
25. Michaels, A.J., et al., Outcome from injury: general health, work status, and satisfaction 12 months after trauma. J Trauma, 2000. 48(5): p. 841-8; discussion 848-50.
26. Michaels, A.J., et al., Posttraumatic stress disorder in injured adults: etiology by path analysis. J Trauma, 1999. 47(5): p. 867-73.
27. Richmond, T.S. and D. Kauder, Predictors of psychological distress following serious injury. J Trauma Stress, 2000. 13(4): p. 681-92.
28. Schnyder, U., et al., Incidence and prediction of posttraumatic stress disorder symptoms in severely injured accident victims. Am J Psychiatry, 2001. 158(4): p. 594-9.
29. Schnyder, U., et al., Posttraumatic stress disorder following accidental injury: rule or exception in Switzerland? Psychother Psychosom, 2008. 77(2): p. 111-8.
30. Shalev, A.Y., et al., Predictors of PTSD in injured trauma survivors: a prospective study. Am J Psychiatry, 1996. 153(2): p. 219-25.
31. Zatzick, D., et al., A national US study of posttraumatic stress disorder, depression, and work and functional outcomes after hospitalization for traumatic injury. Ann Surg, 2008. 248(3): p. 429-37.
32. Zatzick, D.F., et al., Association between posttraumatic stress and depressive symptoms and functional outcomes in adolescents followed up longitudinally after injury hospitalization. Arch Pediatr Adolesc Med, 2008. 162(7): p. 642-8.
33. Zatzick, D.F., et al., Posttraumatic stress, problem drinking, and functional outcomes after injury. Arch Surg, 2002. 137(2): p. 200-5.
34. Zatzick, D.F., et al., Posttraumatic stress disorder and functioning and quality of life outcomes in a nationally representative sample of male Vietnam veterans. Am J Psychiatry, 1997. 154(12): p. 1690-5.
35. Mayou, R., S. Tyndel, and B. Bryant, Long-term outcome of motor vehicle accident injury. Psychosom Med, 1997. 59(6): p. 578-84.
36. Mayou, R. and B. Bryant, Consequences of road traffic accidents for different types of road user. Injury, 2003. 34(3): p. 197-202.
37. Haagsma, J.A., et al., WP5 Injury disability indicators. 2009, Department of Public Health, Erasmus Medical Center: Rotterdam.
38. Bronner, M.B., et al., A community-based survey of posttraumatic stress disorder in the Netherlands. J Trauma Stress, 2009. 22(1): p. 74-8.
39. de Vries, G.J. and M. Olff, The lifetime prevalence of traumatic events and posttraumatic stress disorder in the Netherlands. J Trauma Stress, 2009. 22(4): p. 259-67.
40. Haagsma, J.A., et al., Alternative approaches to derive disability weights in injuries: do they make a difference? Qual Life Res, 2009. 18: p. 657-665.
41. Holtslag, H.R., et al., Individual and population burdens of major trauma in the Netherlands. Bull World Health Organ, 2008. 86(2): p. 111-7.
42. Darves-Bornoz, J.M., et al., Main traumatic events in Europe: PTSD in the European study of the epidemiology of mental disorders survey. J Trauma Stress, 2008. 21(5): p. 455-62.
43. Matsuoka, Y., et al., Incidence and prediction of psychiatric morbidity after a motor vehicle accident in Japan: the Tachikawa Cohort of Motor Vehicle Accident Study. Crit Care Med, 2008. 36(1): p. 74-80.

Part IV

Minimal disease

10

International variation in clinical injury incidence: exploring the performance of indicators based on health care, anatomical and outcome criteria

Suzanne Polinder, Willem Jan Meerding, Ronan A. Lyons, Juanita A. Haagsma,
Hidde Toet, Eleni Th. Petridou, Saakje Mulder, Ed F. van Beeck,
and the EUROCOST reference group
Accident Analysis & Prevention 2008; 40: 182-191

Abstract

Objective To analyse international variation in clinical injury incidence, and explore the performance of different injury indicators in cross-country comparisons.

Methods Hospital discharge data of seven European countries (Austria, Denmark, Ireland, Netherlands, Norway, England and Wales) were analysed. We tested existing and newly developed indicators based on (a) health care use, (b) anatomical criteria, or (c) expected health outcome: admissions excluding day-cases (a), hospital stay 4+ (a) and 7+ days (a), (serious) long-bone fractures (b), selected radiological verifiable fractures 'SRVFs' (b), and indicators based on international (Global Burden of Disease) and Dutch disability weights (c). Assessment criteria were reduction in incidence variation and length of stay in hospital, and the association between incidence and mortality rates.

Results Indicators based on health care use led to increased variation in incidence rates. Long bone fractures and SRVFs, and both indicators based on injuries with moderate to high disability showed similar variation in clinical incidence compared to the crude rates, smaller variation in median length of stay in hospital and a good association with mortality rates.

Conclusion No perfect or near perfect indicators of clinical injury incidence exist. For international comparisons, indicators based on disability weights, SRVFs and long bone fractures may be sensible indicators to use, in the absence of a direct measure of anatomical severity.

Introduction

Injuries are a major cause of morbidity and mortality in the industrialized world [1, 2]. Valid data on the occurrence of injuries are therefore needed. Differences in incidence by time periods, person characteristics and geographic region, and the association between incidence and exposure to risks and safety practices may help policy makers to identify controllable determinants and develop preventive interventions. Internationally, disease and injury incidence is often measured with the help of hospital-based data systems. These data can potentially be used as a measure of incidence to examine international differences in 'clinical injury' (i.e. hospital admitted injury) occurrence. However, the observed variation in clinical injury incidence can be disturbed by differences in admission policy and registration practices between countries. Differences in surveillance data on clinical injuries are generally caused by variation in the definition and classification of injury, proportions of missing data, and the handling of readmissions and day case admissions. The differences in health care policy are mainly related to international differences in hospital admission policy and primary care givers that act as gatekeepers. This is due to the heterogeneity of injuries admitted to hospital, which may range from mild to very severe trauma. This artificial international variation in clinical injury incidence due to variation in health care policy and registration bias could be minimized using a severity threshold for the case definition [3]. Several indicators have been proposed recently [3-8], but a debate remains around the appropriateness of indicators for measuring the occurrence of injuries using administrative datasets. The proposed indicators have not previously been subject to a comparative analysis using international data.

This chapter aims to contribute to the development and validation of injury indicators by applying previously proposed and newly developed injury indicators to clinical injury incidence data across seven European countries. We used predefined assessment criteria to explore whether variation in clinical injury incidence between countries could be real (i.e. based on variation in injury occurrence) rather than artificial (i.e. based on variation in registration and health care practice).

We have addressed the following questions: (a) does clinical injury incidence vary between countries, using different indicators, based on either health care, anatomical or outcome criteria?, (b) which indicator results in the lowest variation in clinical incidence, and the lowest variation in length of stay in hospital?, and c) which indicator results in the best association between clinical incidence and mortality rates?

Methods

Data sources

Within the framework of an international project (EUROCOST), we collected hospital admission data for the year 1999 to produce estimates of the incidence and costs of injuries [9]. Data from the following European countries were analyzed: Austria, Denmark, Ireland, Netherlands, Norway, UK England, and UK Wales. These seven countries were selected because they could provide Hospital Discharge Register (HDR) data as primary data sources to estimate injury incidence of admitted patients (= clinical injury incidence) for intentional and unintentional injuries. The HDR includes general and university hospitals, and private and public hospitals. The data systems that were used are presented in table 10.1 All countries had complete HDR system coverage, except Ireland where 95% of the hospitals contribute to the HDR. The excluded 5% refers to private hospitals. The clinical incidence data from most countries included about 20% day cases, but were lower in Austria (5%), Norway (12%) and Denmark (15%).
For most participating countries, no distinction could be made between first admissions and readmissions. Readmissions are therefore included in all the analyses. For the countries with available information (Austria 0.7%; the Netherlands 2.6%; Norway 8.6%), readmissions are mainly caused by hip fractures, skull-brain injuries and poisoning, resulting in an overestimate of injury incidence for these injury types for most countries. Comparing the health insurance between the countries, in Austria and the Netherlands a mixture of social and private insurances is used. In all other countries health care is financed from tax money.

Injury definition and classification

We applied a uniform case definition to the hospital admission data in accordance with previous studies [5, 10, 11]. Only patients with an injury as principal diagnosis were included (ICD-9 800-995, ICD-10 S00-T78) [12]. We excluded patients with an injury due to medical adverse events (ICD-9 E996-999, ICD-10 T80-T88), early complications (ICD-9 E958, ICD-10 T79) or late effects of injury (ICD-9 905-909, ICD-10 T90-T98). Injuries were classified by location and type into 39 groups after consultations with experts in traumatology, orthopaedics, and rehabilitation, to represent groups of patients that have relatively homogeneous health care consumption [13].

Calculations

Incidence rates were standardized by age (18 groups) and sex, using the direct method of standardization. The total population of the participating countries was used as the reference population. The results are also shown as comparative morbidity figures (CMF). The CMF is calculated by dividing the expected number of injured persons incidence rate (age-specific incidence rates of the country multiplied with population numbers of the

standard population) by the observed number of injured persons of the country. A coefficient of variation was calculated as the ratio of the unweighted standard deviation to the unweighted mean of the outcome measure.

Table 10.1 Description of the hospital data system used and type of insurance per country

Country	Name data system	Coverage (%)	Registered cases	ICD-version	Day cases (%)	Read-missions (%)	Type of insurance
Austria	Hospital Discharge Register	100	187,225	ICD-9	5.1	0.7	Social/ private[2]
Denmark	National Discharge Register	100	99,618	ICD-10	15.3	n.a.[1]	Tax[3]
Ireland	Hospital Inpatient Enquiry	95	58,196	ICD-9	20.8	n.a.[1]	Tax[3]
Netherlands	Dutch Information System	100	102,768	ICD-9	21.4	2.6	Social/ private[2]
Norway	National Injury Register	100	66,962	ICD-10	12	8.6	Tax[3]
England	Hospital Discharge Register	100	632,179	ICD-10	23.1	n.a.[1]	Tax[3]
Wales	Hospital Discharge Register	100	48,266	ICD-10	24.2	n.a.[1]	Tax[3]

[1] n.a. = not available

[2] Mixture of social and private insurances

[3] Financing health care from social taxes (Beveridge-system)

Severity thresholds/case definitions

We applied several indicators for clinical incidence of injury, based on (a) health care criteria (health care use) (b) anatomical criteria (location and type of injury) and (c) outcome criteria (expected level of disability). In Appendix 3.A an overview is given of the inclusions of injury groups per injury indicator.

We selected indicators that were applicable to the national administrative hospital databases of the participating countries. These databases provided no opportunities to calculate established injury severity measures, such as AIS/ISS or ICISS. Our explorative analysis therefore focuses on proxy measures that could be used in the absence of a direct measure of anatomical injury severity.

(a) Health care criteria

1. Exclusion of day cases (new indicator). In the literature [5], it is argued that when day cases are not excluded, biases in estimates of clinical incidence will occur. Day cases and inpatients with length of stay of zero were both classified as a day patient. We tested the performance of a simple new indicator (i.e. all hospital admissions minus day cases).
2. Injuries requiring a hospital stay of 4 days or more ('UK White Paper') [14].
3. Injuries requiring a hospital stay of 7 days or more (new indicator). Our hypothesis was that variation in hospital inpatient rates reduces using higher length of stay thresholds (4+ and 7+ days).

(b) Anatomical criteria

4. Selected radiological verifiable fractures 'SRVFs' (new indicator) [15]. SRVFs include fractures that have radiological verification according to standard practice, and use of this indicator should reduce variation related to the process of data collection and coding [15]. The SRVFs indicator has in fact been defined for the comparison of incidence rates of fractures treated at the Emergency Department rather than the incidence of hospitalizations. However, since hospital admission rates due to SRVFs are high, it seems reasonable to examine the performance of this indicator for the comparability of clinical incidence. At a ProFaNe (Prevention of Falls Network Europe) meeting in 2003 this indicator was advocated for the measurement and evaluation of falls related injury in older people [16].

5/6. Serious long-bone fractures (Cryer) [7] and long-bone fractures (new indicator).
Long-bone fractures were defined as cases admitted to hospital with primary diagnosis of fracture of the femur, or fractures of other long bones of the upper and lower limbs that require an operative procedure. The need for an operative procedure has been proposed as a proxy measure of abbreviated injury scale (AIS), severity score of three or greater) [6]. We used length of stay in hospital as a proxy for injury severity, since no AIS and /or Injury Severity Score (ISS) information was available in our data. McClure suggested that an indicator using 9+ days stay was optimal in terms of sensitivity and specificity when using a threshold of ISS> 15 as the definition of a case of serious injury [6]. We followed this suggestion and defined 'serious long bone fractures' as long bone fractures with hospital duration of nine days or more. In addition, we tested the performance of the more simple potential new indicator 'long bone fractures', which made no prior exclusions by length of stay.

(c) Outcome criteria

We have developed two new indicators based on injury outcome, represented in disability weights. A disability weight reflects the impact of a certain type of injury with a value ranging from 1, indicating best possible health state, and 0, indicating worst possible

health state. Subsequently, the weight of the health state is assessed using a pre-existing set of preference weights, which rest on preference data of a population.

7. Injuries with moderate to high disability weight (Global Burden of Disease) [17]. The Global Burden of Disease (GBD) study [1, 17] provided disability weights by nature of injury using the person-trade-off method. Participants were asked to trade-off absolute numbers of healthy individuals against absolute numbers of individuals in a given suboptimal health state. Disability weights provide a summary measure of the severity of the expected health consequences of diseases and injury. All disability weights can be found in Murray [17]. This indicator includes injuries with a moderate to high disability weight according to the GBD study, defined as a disability weight for treated forms of sequelae higher than 0.20.
8. Injuries with moderate to high disability weight (Dutch study). A Dutch study provided disability weights by type of injury representing the severity of the health state following injury in survivors. These weights were derived using the time-trade-off method (TTO) and the visual analogue scale (VAS). In the TTO, participants were asked how much time they were willing to trade in order to be restored from the presented health state to full health. With the VAS the participants scored the health states on a vertical thermometer graded from 0 (worst imaginable health state) to 100 (best imaginable health state). Subsequently, the VAS and TTO values were transformed into disability weights. The disability weights range from 0.005 for superficial injury to 0.638 for spinal cord injury. The assumption is made that the Dutch disability weights were applicable for the other participating countries. We selected injuries with a moderate to high disability weight, defined as a disability weight higher than 0.10.

Assessment criteria

We assessed each of these indicators according to the following criteria:

1. Reduction in the international variation in clinical incidence (as measured by the coefficient of variation).
2. Reduction in the international variation in median length of hospital stay. This assessment criterion was not applied to the health care criteria.
3. Comparable ranking of countries according to the clinical incidence rates and according to mortality. The injury related mortality rates are based on the World Health Organization Statistics [18].

We hypothesized that when the case definition was restricted to injury groups with a severity threshold indicating a necessity for hospital admission that international variation in clinical incidence and hospital length of stay would reduce. Also, assuming more or less similar quality of care and case fatality rates in European countries and a distribution of exposures producing a similar distribution of injury severity then, clinical incidence rates would be more strongly associated with mortality rates.

Results

Variation in clinical incidence

In table 10.2, the clinical incidences and comparative morbidity figures are shown for the crude data and the injury indicators based on health care criteria. The age-standardized clinical incidence varied substantially using crude data, ranging from 6.6 to 22.9 per 1,000 person years. Austria ranked highest, followed by Denmark, Wales, Ireland, Norway, England, and the Netherlands.
Excluding day cases in the analysis resulted in larger variation of injury incidence between the countries, which is shown by and increased coefficient of variation. When patients with a length of stay shorter than four or seven days were excluded, this also resulted in an increased variation of clinical incidence. For Wales, Ireland, Norway, England, and the Netherlands the clinical incidence rate for long stay admitted injury patients (hospital stay > 7 days) varied between 2.0 and 2.7 per 1,000 inhabitants. Only in Denmark (4.6) and Austria (6.6) substantially more patients were admitted after an injury for longer than a week.

Table 10.2 Clinical injury incidence (per 1,000) and CMF for injury indicators based on health care criteria (assessment criteria 1)

	Crude data		Excluding day cases		Hospital stay > 4 days		Hospital stay > 7 days	
Country	Incidence	CMF	Incidence	CMF	Incidence	CMF	Incidence	CMF
Austria	22.9	1.5	21.7	1.8	11.1	2.4	6.6	2.2
Denmark	18.1	1.2	15.4	1.3	6.9	1.5	4.6	1.5
Wales	16.0	1.1	12.3	1.0	4.1	0.9	2.7	0.9
Ireland	15.5	1.0	12.5	1.0	3.6	0.8	2.8	0.9
Norway	14.7	1.0	12.9	1.1	4.8	1.0	2.9	0.9
England	11.9	0.8	9.1	0.8	3.1	0.7	2.1	0.7
Netherlands	6.6	0.4	5.2	0.4	2.8	0.6	2.0	0.7
Coefficient of variation	0.34		0.40		0.56		0.48	

In table 10.3, the clinical incidences and comparative morbidity figures are shown for the the injury indicators based on anatomical and outcome criteria. All indicators based on anatomical criteria resulted in reduced or similar variation in clinical incidence. For selected radiological verifiable fractures (SRVF) and long bone fractures an almost similar and consistent pattern of international variation in clinical incidence was observed. Higher than average incidence rates were found for Austria, Denmark, Norway and Ireland, whereas incidence rates on or below the average were observed for England, Wales, and the Netherlands.

The indicators based on expected levels of disability resulted in a slightly increased variation in clinical incidence for the GBD weighting scale, whereas the Dutch disability weight showed a somewhat lower variation in clinical injury incidence than the crude data. For the disability indicators the same dichotomy arises as was found for the anatomical criteria, with 11/2-2 times higher incidence rates for Austria, Denmark and Norway, and Ireland compared with Wales, England and the Netherlands.

Table 10.3 Clinical injury incidence (per 1,000 inhabitants) and CMF for the indicators based on anatomical and outcome criteria (assessment criteria 1)

	SRVF		Serious long bone fractures		Long bone fractures		High disability (GBD)		High disability (GBD)	
Country	Inc[1]	CMF	Inc[1]	CMF	Inc[1]	CMF	Inc[1]	CMF	Inc[1]	CMF
Austria	5.5	1.9	2.6	1.8	4.7	1.7	4.1	1.7	4.7	1.7
Denmark	5.2	1.7	2.3	1.6	4.6	1.6	4.4	1.8	4.6	1.6
Wales	2.2	0.7	1.2	0.8	2.5	0.9	2.0	0.8	2.5	0.9
Ireland	3.8	1.3	1.4	0.9	3.7	1.3	2.9	1.1	3.7	1.3
Norway	4.7	1.6	1.7	1.2	3.9	1.4	3.9	1.6	3.9	1.4
England	3.0	1.0	1.4	1.0	1.8	0.6	2.6	1.0	1.8	0.6
Netherlands	2.6	0.9	1.3	0.9	2.1	0.7	2.1	0.9	2.1	0.7
Coefficient of variation	0.34		0.32		0.36		0.32		0.36	

[1] Inc=Incidence

Variation in length of stay

The indicators based on anatomical criteria resulted in a reduced variation in median length of stay in hospital as opposed to the crude data (table 10.4). The median length of stay of patients admitted with serious long bone fractures is quite comparable across countries.
Injuries with a moderate to high disability weight showed less variable hospital durations than the total population on both indicators.

Table 10.4 Median length of stay (days) by injury indicators based on anatomical and outcome indicators (assessment criteria 2)

	Crude data	SRVF	Serious long bone fractures	Long bone fractures	High disability (GBD)	High disability (IBIS)
Austria	4	9	16	10	9	10
Denmark	2	6	15	7	6	7
Ireland	2	4	13	4	4	4
Netherlands	2	8	16	9	10	11
Norway	2	5	12	5	5	5
England	1	7	14	8	7	7
Wales	1	10	15	10	8	9
Coefficient of variation	0.50	0.31	0.10	0.31	0.31	0.34

Ranking order clinical incidence versus mortality

Comparing the clinical incidence rates of indicator 3 (hospital stay > 7 days) with the mortality ranking, Austria and Denmark demonstrated the highest mortality rates (see figure 10.1). Mortality in Norway, however, was almost twice as high as in The Netherlands, England and Wales, whereas the clinical incidence rates of indicator 3 (hospital stay > 7 days) were low for all these four countries. Contrary to the crude data, indicators based on anatomical criteria showed reasonable (serious long bone fractures) to good (SRVF and long bone fractures) associations with mortality rates. For SRVF and long bone fractures the clinical incidence of Norway ranked consistently higher and Wales ranked consistently lower compared to the crude incidence ranking, which corresponds with the mortality ranking. The coefficient
of variation of the indicators based on anatomical criteria (0.31–0.34) was almost equal to the coefficient of variation of the mortality rates (0.29: data not shown). Using injury indicators based on disability, Norway ranked consistently higher and Wales ranked consistently lower compared to the crude incidence ranking. Both disability indicators show a quite comparable clinical incidence pattern compared to the mortality ranking.

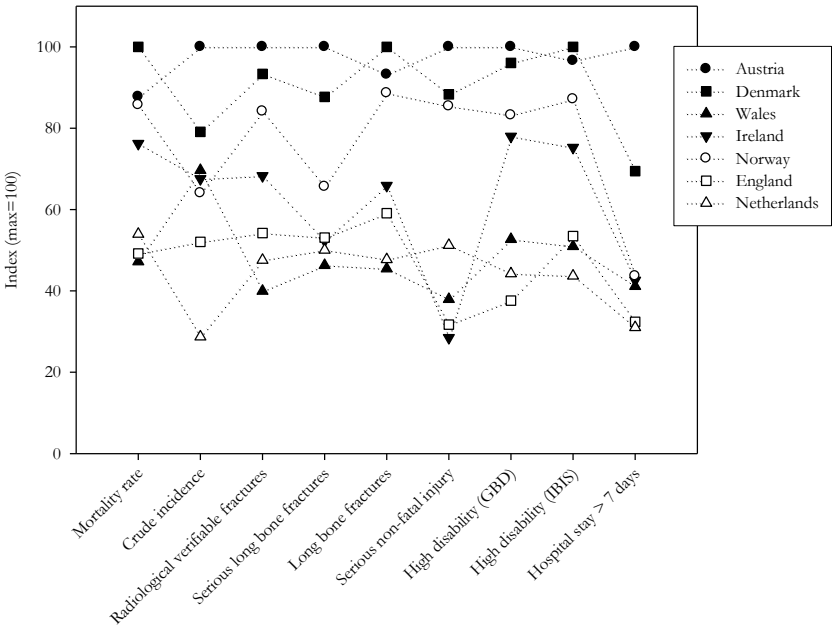

Figure 10.1 Country ranking according to the clinical incidence rates and mortality by injury indicators (assessment criteria 3)

Discussion

This chapter described four previously proposed and four new injury indicators and explored their validity with use of three assessment criteria. Indicators based on health care criteria led to increased international variation in incidence rates and a different pattern of international variation compared to mortality rates. The anatomical indicators (serious) long bone fractures and SRVFs, as well as both indicators based on expected disability did not result in a substantial reduction of the variation in clinical incidence. However, they were associated with reduced variation in median length of stay in hospital and showed fair to good associations with mortality rates. Whatever measure is used, international differences in clinical incidence of injuries of moderate to high severity remain considerable, and are higher than international differences in injury mortality.

Indicators based on health care criteria

In earlier research, McClure introduced 'injuries that result in hospitalization' as an indicator [6]. However, indicators based on hospital admission of any severity or diagnosis are biased and unstable [4]. Indicators based on length of stay in hospital have the disadvantage that they are sensitive to changes due to service factors [4]. Furthermore, since the proportion of day cases was the lowest in the country with the highest incidence (Austria) the international variation of the clinical incidence rate increased when day cases were excluded. In this explorative study, the indicators based on health care criteria led to increased variation in clinical incidence rates, which supports earlier recommendations to avoid their use.

Indicators based on anatomical criteria

Since differences in performance between serious long bone fractures and long bone fractures were not substantial, we prefer the less complex indicator, i.e. long bone fractures. However, the main disadvantage of (serious) long bone fracture is that it is only useful as an indicator for serious extremity injury, whereas it fails to consider important mortality and disability causing injuries, such as skull-brain, and spinal cord injury. The indicator SRVFs was developed as an indicator for injuries treated at the ED (and subsequently admitted or not). It therefore includes injuries that are not necessarily severe, and would not be treated as inpatients in most countries. Examples are fractures of wrist and lower leg. However, both SRVFs and 'long bone fractures' satisfied two of our assessment criteria and are therefore recommended for further research and application if direct measures of anatomical severity cannot be obtained.

Indicators based on outcome criteria

We introduced new indicators based on disability weights and showed them to be a possible alternative for indicators based on anatomical criteria. For injuries with a moderate to high disability weight the variation in clinical incidence and the variation in hospital duration were similar to the variation using crude data. However, variation in length of stay in hospital was reduced, and these indicators showed a similar pattern of variation as with mortality rates. Further refinement and testing of these new indicators is therefore recommended.

Strengths and weaknesses of the study

To our knowledge, this is the first study in which existing and newly developed injury indicators were all together empirically tested with hospital admission data of seven European countries. This demonstrates that it is feasible to apply these indicators to routinely collected data. Nevertheless, our explorative study has several limitations and further work is necessary to validate and improve injury indicators.

The use of routinely available administrative national data systems bears some weaknesses. First of all, the quality of recording, coding and classifying injuries may differ between countries. For example, several of our comparisons used specific selections of injury diagnoses, assuming comparability of injury type and severity between countries for those selections. Differential misclassification within the administrative databases between countries may have occurred and affected the results of our comparisons of both clinical incidence rates and median length of stay in hospital. International differences in data quality can compromise attempts to standardize injury indicators but can only be identified with the help of in-dept research. Therefore, additional validation studies are recommended, with samples of hospitalized injury data from several countries being selected and judged by an expert "gold standard" panel.

Our comparison of clinical incidence rates with mortality rates also has limitations, since the case fatality rate of injury may vary between countries. The absence of empirical data makes this assumption difficult to prove or disprove. However, by focusing on a rather homogeneous set of countries in Europe, we think this problem has probably largely been avoided in our study. Nevertheless, additional study of variation in clinical case fatality rates for specific selections of injuries, which could not be conducted with the available data, is recommended. In our explorative analysis of indicators we used several proxies for severity, duration of hospitalization, type of injury, and expected level of disability. In the data available to us, it was not possible to define an injury severity threshold using a metric such as the Abbreviated Injury Scale (AIS). The availability of adequate data on direct anatomical measures of injury severity in (national) databases could further enhance international data comparability and would make the injury indicators more useful for policy making, especially with regard to injuries of higher severity levels. With this purpose in mind Cryer and colleagues developed the 'age-standardized serious non-fatal injury' indicator [10]. This indicator is based on ICISS scores, i.e. it is a threat-to-life severity scale based on crude in-hospital mortality rates by ICD-code [19]. In principle, this indicator is applicable to ICD-based administrative databases. However, since the participating countries in our study did not provide individually specified ICD-codes we could not exactly distinguish ICD-10 and ICD-9 principle diagnoses by the ICISS severity threshold and could not test the performance of this indicator against our assessment criteria. Comparing the crude incidence data, Austria has a very high clinical incidence for all injury groups. Austria has also a relatively high clinical incidence of low severity injuries (superficial and remaining injuries, mainly dislocation/sprain/strain knee, fractures arm/ankle, and open wounds), which indicates a low admission threshold for this country. At the opposite, it seems that the Netherlands has a high admission threshold, indicated by their high mean length of stay (9.3 days compared to a mean of 6.4 days for all countries).

Recommendations

Our study must be interpreted as an exploration of the performance of injury indicators in an international perspective. For the further validation of specific injury indicators we recommend using larger samples of injury patients from more regions within countries, to increase the power and validity of the results.
Nevertheless, the observed effects for various case selection criteria serve as a warning for undertaking international comparisons of clinical incidence of injury. Our findings question the feasibility of comparing comprehensive and therefore heterogeneous injury populations in trend analyses or international comparisons. In particular hospital admission rates for minor injuries are influenced by large differences in registration practice and health care policy. In order to reduce these methodological problems it is better to focus on the more severe injuries. However, the choice of indicator of injury occurrence has a substantial influence on international comparisons and no perfect or near perfect indicators have yet been developed. Indicators based on disability weights or on anatomical criteria satisfied two of our three assessment criteria. If we assume that indicators should focus attention on important problems [4], then indicators based on disability weights (GBD and Dutch weights) may be sensible indicators to use, in the absence of a direct measure of anatomical severity. The long bone fractures indicator performs quite well, but has the disadvantage that it does not include some important injuries with a high burden in terms of mortality and disability. The same argument applies to the SRVFs indicator, which was not developed for this purpose but for counting ED-treated injury patients [15]. The choice of indicator will depend on the specific research or policy question addressed. For international comparison of the clinical incidence of home and leisure injuries, the long bone fractures indicator might satisfy, because the impact on health among this group is largely dominated by fractures, particularly in the elderly. For international comparison of the clinical incidence of traffic injuries, a combined indicator based on mortality and disability weights might be preferable, since these indicators also include injuries to the head, spine and internal organs, which are frequently fatal or very disabling for a substantial part of hospital-admitted traffic victims.

We recommend avoiding the naïve use of hospital-based data systems for international comparisons and national trend analyses. The risk of measuring artificial instead of real differences in injury incidence between countries, patient groups, or time periods is large. The methodological considerations and developments recommended in this chapter should be further refined and tested before they can be used as tools for those who need to compare injury incidence data.
Many of the injury indicators of non-fatal injury analysed in this chapter are based on moderately high thresholds of severity. There is merit in considering the development of non-fatal injury indicators that capture less severe cases, but which exclude minor/superficial injury. It should be recognized, that this is not likely to be possible

using hospital inpatient data alone [10], since many moderately severe injuries are treated in outpatient settings without recourse to admission to hospital. Further research requires studies utilizing linkage of individual level outpatient (and emergency department) and inpatient data with high quality coding and preferably severity measures incorporating both threat to life and threat to disability scales. Data from several ongoing and planned prospective burden of injuries studies should help answer these questions.

References

1. Murray, C.J. and A.D. Lopez, Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. The Lancet, 1997. 349(9063): p. 1436-42.
2. Morrison, A., D.H. Stone, and EURORISC working group, Injury mortality in the European Union 1984-1993. Eur J Public Health, 2000. 10: p. 201-7.
3. Cryer, C., et al., Injury outcome indicators: the development of a validation tool. Inj Prev, 2005. 11: p. 53-57.
4. Cryer, C., et al., Measure for measure: the quest for valid indicators of non-fatal injury incidence. Public Health, 2002. 116(5): p. 257-62.
5. Langley, J.D., G.S. Davie, and J.C. Simpson, Quality of hospital discharge data for injury prevention. Inj Prev, 2007. 13(1): p. 42-4.
6. McClure, R.J., et al., Appropriate indicators for injury control? Public Health, 2002. 116(5): p. 252-6.
7. Cryer, P.C., et al., How can we reliably measure the occurrence of non-fatal injury? Int J Consum Prod Saf, 1999. 6(4): p. 183-91.
8. Cryer, C. and J.D. Langley, Developing valid indicators of injury incidence for "all injury". Inj Prev, 2006. 12: p. 202-207.
9. Polinder, S., et al., Cost estimation of injury related hospital admissions in ten European countries. Journal of Trauma, 2005. 59(6): p. 1283-1291.
10. Cryer, C., J. Langley, and S. Stephenson, Developing valid injury outcome indicators: a report for the New Zealand injury prevention strategy, O.R.O. 049, Editor. 2004, Occasional Report OR 049: Dunedin, New Zealand.
11. Moustaki, M., M. Lariou, and E. Petridou, Cross country variation of fractures in the childhood population. Is the origin biological or "accidental"? Inj Prev, 2001. 7(1): p. 77.
12. WHO, International Classification of Diseases; 9th edition. 1997, World Health Organisation: Geneva.
13. Meerding, W.J., et al., A surveillance based assessment of medical costs of injury in Europe: phase 1. 2002, Consumer Safety Institute: Amsterdam.
14. Secretary of State for Health, Saving Lives: Our Healthier Nation, in The Stationery Office, T.S. Office, Editor. 1999: London.
15. Lyons, R.A., et al., Methodological issues in comparing injury incidence across countries. Int J Inj Contr Saf Promot, 2006. 13(2): p. 63-70.
16. ProFaNE. http://www.profane.eu.org. [cited; Available from: http://www.profane.eu.org.
17. Murray, C.J.L. and A.D.e. Lopez, The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. 1996, Cambridge: Harvard University Press.
18. WHO, World Health Organization mortality database. Available from: http://www3.who.int/whosis/menu.cfm?path=whosis,mort&language=english.
19. Rutledge, R., et al., Comparison of the Injury Severity Score and ICD-9 diagnosis codes as predictors of outcome in injury: analysis of 44,032 patients. J Trauma, 1997. 42(3): p. 477-87; discussion 487-9.

Appendix 3.A Inclusion of injury groups by injury indicator

	Health care criteria		Anatomical criteria		Outcome criteria	
39 Injury groups	Exclusion day cases	Hospital stay > 4/7 days	Radiological verifiable fractures	Long bone fractures	High disability (IBIS)	High disability (GBD)
Head						
1. Concussion	X	X				
2. Skull-brain injury	X	X			X	X
3. Open wound head	X	X				
Face						
4. Eye injury	X	X				X
5. Fracture facial bones	X	X				X
6. Open wound face	X	X				
Vertebrae / Spine						
7. Vertebral column fractures dislocations/sprain/strain	X	X				X
8. Whiplash, neck sprain, distortion of cerval spine	X	X				
9. Spinal cord injury	X	X			X	X
Abdomen / Thorax						
10. Internal organ injury	X	X			X	X
11. Fracture rib / sternum	X	X				
Upper extremity						
12. Fracture of clavicle / scapula	X	X				
13. Fracture of upper arm	X	X	X	X		
14. Fracture of elbow / forearm	X	X	X	X		
15. Fracture of wrist (incl. carpal bones)	X	X	X			
16. Fracture of hand/fingers	X	X				
17. Dislocation / sprain / strain shoulder / elbow	X	X				
18. Dislocation / sprain / strain wrist/hand/ fingers	X	X				
19. Injury of nerves	X	X				
20. Complex soft tissue injury upper extremity	X	X			X	
Lower extremity						
21. Fracture of pelvis	X	X	X		X	X
22. Fracture of hip	X	X	X	X	X	X

39 Injury groups	Health care criteria		Anatomical criteria		Outcome criteria	
	Exclusion day cases	Hospital stay > 4/7 days	Radiological verifiable fractures	Long bone fractures	High disability (IBIS)	High disability (GBD)
23. Fracture of femur shaft	X	X	X	X	X	X
24. Fracture knee/lower leg	X	X	X	X		
25. Fracture of ankle	X	X	X			
26. Fracture of foot	X	X				
27. Dislocation / sprain / strain of knee	X	X				
28. Dislocation / sprain / strain of ankle / foot	X	X				
29. Dislocation / sprain / strain of hip	X	X			X	
30. Injury of nerves	X	X				
31. Complex soft tissue injury lower extremity	X	X			X	X
Minor external						
32. Superficial injury (incl. contusions)	X	X				
33. Open wounds	X	X				
34. Burns	X	X				
35. Poisoning	X	X				
36. Multitrauma	X	X				
Other injuries						
37. Foreign body	X	X				
38. No injury after examination	X	X				
39. Other and unspecified injury	X	X				

11

Disability adjusted life years and minimal disease: application of a preference-based relevance criterion to rank enteric pathogens

Juanita A. Haagsma, Arie H. Havelaar, M.F. (Bas) Janssen, Gouke J. Bonsel
Population Health Metrics 2008; 6:7-15.

Abstract

Background Burden of disease estimates, which combine mortality and morbidity into a single measure, are used increasingly for priority setting in disease control, prevention and surveillance. However, because there is no exclusion criterion for highly prevalent minimal disease in burden of disease studies its application may be restricted. The aim of this study was to apply a newly developed relevance criterion based on preferences of a population panel, and to compare burden of disease estimates of five foodborne pathogens calculated with and without application of this criterion.

Methods Preferences for twenty health states associated with foodborne disease were obtained from a population panel (n = 107) with the Visual Analogue Scale and the Time Trade-off (TTO) technique. The TTO preferences were used to derive the relevance criterion: if at least 50% of a panel of judges is willing to trade-off time in order to be restored to full health the health state is regarded as relevant, i.e. TTO median is greater than 0. Subsequently, the burden of disease of each of the five foodborne pathogens was calculated both with and without the relevance criterion.

Results The panel ranked the health states consistently. Of the twenty health states, three did not meet the relevance criterion. Application of the preference-based relevance criterion reduced the burden of disease estimate of all five foodborne pathogens. The reduction was especially significant for norovirus and rotavirus, decreasing with 94% and 78% respectively.

Conclusions Individual preferences elicited with the TTO from a population panel can be used to empirically derive a relevance criterion for burden of disease estimates. Application of this preference-based relevance criterion results in considerable changes in ranking of foodborne pathogens.

Introduction

Since the application of the concept in 1993, the Disability Adjusted Life Year (DALY) is used increasingly for priority setting in disease control, prevention and surveillance [1]. The DALY is a health gap measure that aggregates mortality, acute morbidity and disability into a single index [2]. In order to combine information on mortality and morbidity, firstly the years lost due to premature mortality (YLL), and secondly the years lived with disability (YLD) have to be estimated. The latter result from a computational procedure that combines duration and severity of both acute disease and disability from sequelae.

An essential component of YLD computation is the disability weight. A disability weight is a scaling factor assigned to living with disability that ranges from 0 (best possible health state) to 1 (worst possible health state or equating death) [3]. This value reflects the impact of the disability on the health-related quality of life and is commonly based on the preferences of a panel of judges [4].

To arrive at YLDs, the disability weights have to be combined with incidence and duration data. In some burden of disease studies hospital admissions and Emergency Department treatments are used as data source in this regard [5, 6]. Other studies use data from General Practitioner registries or population health surveys [7, 8]. Although the latter approach yields otherwise lacking incidence data which are vital to YLD calculations, such registries and surveys have low response thresholds, implying that cases of minimal disease are also included.

Minimal disease consists of temporary health states that have an anticipated and observed minor impact on individual health-related quality of life. However, due to high prevalence of most minimal disease, collectively they may account for a large number of YLDs in the aggregate, and may therefore get policy priority above severe, but less frequently occurring diseases. As a result, the application of burden of disease estimates in prioritization discussion may be obfuscated. This can be overcome by including a criterion for relevant disease. A prerequisite of this relevance criterion is that it is able to distinguish 'experienced' minimal disease from relevant disease unambiguously, yet it should also allow relevant mild disease to be included in burden of disease estimates in order to avoid incomplete estimates of diseases characterized by heterogeneous levels of severity.

Other than on health care use, the criterion to distinguish trivial diseases from relevant diseases may be based on health outcome; anatomical characteristics of the disease, absenteeism (work, school), or on societal preferences that are derived to assess the disability weights necessary for the YLD calculation. To obtain the latter, preferences from a panel of judges, elicited with dedicated preference measurement methods commonly are used [9, 10]. A preference measurement method widely used in these panel studies is the time trade-off (TTO) method. The TTO method requires an individual to give up time in order to be restored from the health state to full health [10]. The more time the participant is willing to offer, the less desirable the health state is compared with

perfect health. In the Dutch Mild Diseases and Ailments study, Bonsel et al. developed a relevance criterion based on the TTO preferences to distinguish relevant from minimal conditions [11]. The preference-based relevance criterion is met if the median TTO value is greater than 0. i.e. if at least 50% of a panel of judges is willing to trade-off time in order to be restored to full health. If not, the health state is regarded as not relevant and excluded from the burden of disease calculation. The cut-off point of the criterion, where at least 50% of the panel members has to be willing to trade-off time, corresponds to the majority rules principle of most democratic voting systems.
In the area of foodborne disease, often high incidences of infectious disease are observed with most cases leading to full recovery within only a few days. Comprehensive studies on foodborne disease are vulnerable to inadvertently putting too much emphasis on such minimal diseases.
To aim of our study in this context was (1) to derive a relevance criterion for foodborne disease based on the preferences of a population panel, and (2) to compare YLD estimations of five common foodborne pathogens calculated with and without the application of the preference-based relevance criterion.

Methods

Health state description

Five enteric pathogens were selected, namely norovirus and rotavirus, thermophilic *Campylobacter* spp., *Salmonella* spp., and Shiga-toxin producing *Escherichia coli* O157 (STEC O157). We carefully defined the diseases caused by these foodborne pathogens, collected empirical evidence on the associated functional consequences over time and subdivided the diseases into disease severity grades that were presumed to be homogeneous regarding disability, treatment and prognosis. This resulted in 20 health states.
The functional consequences of each of the 20 health states were presented on a vignette. A vignette is a preformatted A4 size sheet that provided the disease label, clinical description and a generic description. For the generic description, we used an extended version of the EQ-5D classification system [12-14]. This classification system describes health with five levels of severity in the dimensions mobility, self-care, usual activities, pain/discomfort, anxiety and depression and cognition [15].
Additionally, the vignette provided a visual representation of the body sites affected by the disease, and described the duration of the disease over 1 year time. The duration of the condition was presented as an annual profile, which describes the course of the health state – the disability profile – over 1 year, allowing assessment of diseases with a rapid course [4, 16]. Conditions with short duration were presented as a patient who in an otherwise healthy year experiences, for instance, the health impact of mild gastroenteritis for the duration of one week; whereas conditions with long-term consequences were

presented as a patient who experiences, for instance, the consequences of Guillain-Barré syndrome throughout the whole year.

Health state valuation

The Visual Analogue Scale (VAS) and the TTO were used to elicit preferences for the 20 health states [17]. The VAS valuation technique requires participants to score the health state on a vertical rating scale graded from 0 (worst imaginable health state) to 100 (best imaginable health state). In the TTO the participants were asked how many days of 1 year in full health they were willing to trade in order to be restored from the presented disease stage to full health. A trade-off of time implies a shorter life expectancy in exchange for full health during the remaining time. The participants were instructed to contemplate the year described by the vignette only and to ignore any prognostic element of what could happen after one year.
All health states were valued independently according to both methods.

Participants and data-collection

The panel participants were randomly selected from a sample of 560 lay people who applied to participate in the Mild Diseases and Ailments Study, conducted in 2003 [16]. For the Mild Diseases and Ailments Study people were recruited from the general public via an advertisement in a newspaper that is freely available throughout The Netherlands. For this study, a random sample of 150 persons was drawn and contacted by mail of which 115 were willing to participate.
Valuation data were collected through a two-step procedure. Firstly, the participants attended a 3 hour panel session, during which they valued 10 vignettes with the VAS and the TTO. The second part of the data-collection consisted of an unsupported postal questionnaire which the participants received at home one week after attending the panel meeting. Apart from the 20 health states related to foodborne disease, the participants valued 24 other health states (core health states inserted to verify reliability among groups, work-related health states and psychiatric health states) totalling the number of questionnaire vignettes to 34. Because of the high total number of vignettes, we chose to develop two versions of the questionnaire. Each version of the questionnaire asked the participant to value 17 vignettes with the VAS and the TTO. The order of the vignettes was randomized and the questionnaire version was randomly assigned to the participants. The participants received 20 euros for attending the panel session and 30 euros for filling in the follow up questionnaire.

Data-analysis

Firstly, we standardized VAS and TTO weights for each individual response using the following formulas:

VAS weight = 1 – (VAS score/100).

TTO weight = TTO score/365.

To establish whether the disease stages were ranked in consistent order, the ranking of a vignette according to the VAS and TTO weights was compared using the Spearman and Pearson's rank correlation coefficients. We performed regression analysis to determine whether socio-demographic variables age, sex, education level and disease experience had independent significant effects on the VAS and TTO weights. To establish the inter-rater reliability, which measures group homogeneity, we defined each of the eight panel meetings as a group and calculated the intra-class correlation (ICC) of these groups.
In order to calculate the YLD, national incidence data on foodborne disease (year 2004) was combined with the disability weight derived from the panel study. YLD was calculated both with and without application of the relevance criterion. The relevance criterion implies that if the median TTO = 0, the disease stage is regarded as trivial and therefore not included in the YLD calculation.
To calculate YLL we used mortality data from Statistics Netherlands. The resulting YLL were combined with YLD in order to calculate the number of DALYs lost due to the five foodborne pathogens, all following standard procedures (no age-weighting). For a more detailed description of the DALY calculation, see Kemmeren et al.[18].

Results

Participants and data-collection

The panel meeting was attended by 107 participants. Each of the 107 participants responded to the questionnaire. On average it took the participants 1 hour and 43 minutes to fill out the questionnaire. The average age of the participants was 51 years and 62% was female. Statistics of the participants are presented in table 11.1.

Table 11.1 Statistics of the population panel

Statistics	Mean (range) or n (%)
Age (in years)	51.2 [21 -79]
Sex	
Male	41 (38%)
Female	66 (62%)
Educational level	
Low	22 (21%)
Middle	43 (40%)
High	41 (38%)
Unknown	1 (1%)
Disease experience	
Yes	55 (51%)
No	51 (48%)
Unknown	1 (1%)

Health state valuations

Table 11.2 presents the mean and median TTO weights for the 20 health states. Mild conditions with short duration, like gastroenteritis, was rated lowest (mean TTO weight 0.01) whereas severe long-term disease, like Guillain-Barré syndrome level F5, was rated highest (mean TTO weight 0.46). Weights increased by level of severity within the diseases; weights attributed to mild reactive arthritis (mean TTO weight 0.02) were lower than moderate reactive arthritis (mean TTO weight 0.12) and severe reactive arthritis (mean TTO weight 0.19) respectively. The percentage of participants that were not willing to trade-off any time decreased by the level of severity within the disease; for moderate haemolytic uremic syndrome 13% of the participants were not willing to trade-off time, whereas for severe haemolytic uremic syndrome this was 0%. The standard deviation of the TTO values was higher in the middle range, which was anticipated given the fixed end points of the scale. Correlation coefficients between VAS and TTO values were high, Pearson's correlation coefficient was 0.92 and Spearman correlation coefficient was 0.95. ICC, which indicates the inter-rater reliability, was 0.99 for the VAS values and 0.97 for the TTO values. No significant effects of age, sex, and disease experience on TTO values were demonstrated. Educational level, however, did have a significant effect on the TTO values of the Guillain-Barré disease stages.

Table 11.2 Mean VAS weights and mean and median TTO weights, per health state

		VAS	TTO		
Health state and length of illness	n	mean	median	mean	%0 [a]
Gastroenteritis					
Mild, 1 day	51	0.036	0	0.002	88
Mild, 5 days	53	0.102	0	0.010	60
Moderate, 10 days	107	0.130	0.005	0.015	26
Severe, 7 days	53	0.231	0.008	0.025	25
Severe, 14 days	51	0.295	0.011	0.041	17
Chronic, 6 months	53	0.368	0.058	0.099	8
Guillain-Barré syndrome [b]					
F1, whole year	51	0.185	0.008	0.044	40
F2, whole year	107	0.420	0.077	0.137	7
F3, whole year	53	0.545	0.153	0.215	2
F4, whole year	51	0.700	0.252	0.367	2
F5, whole year	53	0.722	0.403	0.460	0
Reactive arthritis					
Mild, 1 week	51	0.107	0	0.004	68
Mild, 6 weeks	53	0.197	0.011	0.023	25
Moderate, 6 months	53	0.447	0.058	0.115	8
Severe, 6 months	51	0.503	0.153	0.186	4
Hemolytic Uremic Syndrome					
Moderate, 1 month	53	0.279	0.022	0.056	13
Severe, 1 month	51	0.481	0.038	0.110	0
Renal failure, whole year	51	0.628	0.252	0.328	0
Inflammatory bowel disease					
Crohn disease, 6 months	51	0.347	0.067	0.105	4
Colitis ulcerosa, 6 months	53	0.492	0.115	0.154	7

[a] percentage of participants that were not willing to trade-off any time in order to be restored from the health state.

[b] for a detailed description of the five health states of patients with Guillain-Barré syndrome, see Havelaar et al. [8]

Preference-based relevance criterion

Table 11.2 shows that three health states, namely mild gastroenteritis with a length-of-illness of respectively one day, and one week, and mild reactive arthritis with length-of-illness of one week, had a TTO median of 0. Therefore, these three disease stages did not meet the relevance criterion.
Table 11.3 presents the TTO weights elicited in two previous Dutch health state valuation studies conducted in 2003 and 2005 that had a design similar to the current study [16, 19]. The results show that three health states, namely common cold with a length-of-illness of one week, onychomycosis with a length-of-illness of a whole year, and superficial injury with a length-of-illness of four weeks had a TTO median of 0 in both studies.

Table 11.3 Median TTO weights of 10 health states valued in two preceding panel studies, by health state

	2003			2005		
Health state	n	median	%0 [a]	n	median	%0 [a]
Common cold, 1 week	101	0	85	140	0	81
Cystitis, 2 weeks	102	0.003	47	64	0.005	27
Rhinitis, 17 weeks	102	0.003	38	64	0.019	28
Eczema, whole year	102	0.019	16	64	0.044	11
Gastritis, 4 weeks	101	0.005	37	64	0.019	16
Onychomycosis, whole year	102	0	83	63	0	60
Osteoporosis, whole year	102	0.003	45	142	0.003	36
Otitis, 2 weeks	101	0.003	40	64	0.010	44
Superficial injury, 4 weeks	32	0	53	142	0	72
Open wound, 4 weeks	34	0.005	41	47	0.003	49

[a] %0 = the percentage of participants that were not willing to trade-off any time in order to be restored from the health state

DALY calculation

The burden of disease was calculated for all five foodborne pathogens (see table 11.4). Without application of a relevance criterion most DALYs were lost due to norovirus (2940 DALYs), rotavirus (1327 DALYs) and thermophilic *Campylobacter* spp. (1137 DALYs). Least DALYs were lost due to *Salmonella* spp. (747 DALYs) and STEC O157 (120 DALYs).
Application of the relevance criterion resulted in a burden of disease estimate of 175 DALYs due to norovirus and 287 DALYs due to rotavirus, a decrease of 94% and 84%, respectively. For thermophilic *Campylobacter* spp. the burden of disease reduced by with 24% to 865 DALYs, and for *Salmonella* spp. with 24% to 567 DALYs. With 5%, the decrease in burden of disease was smallest for STEC O157, which reduced from 120 to 114 DALYs.

As a result of the reductions in part of the burden of disease estimates, the ranking of the foodborne pathogens changed. Without a relevance criterion, the ranking according to descending burden of disease was: 1) norovirus, 2) rotavirus, 3) thermophilic *Campylobacter* spp., 4) *Salmonella* spp., and 5) STEC O157. Conversely, when the preference-based relevance criterion was applied the ranking was: 1) thermophilic *Campylobacter* spp., 2) *Salmonella* spp. 3) rotavirus, 4) norovirus, and 5) STEC O157.

Table 11.4 Incidence and disease burden calculated with and without the preference-based relevance criterion (RC), by pathogen

			Without RC		*With RC*	
Pathogen	Incidence	YLL	YLD	DALY	YLD	DALY
Norovirus						
Gastroenteritis	472,000	55	2885	2940	121	175
Rotavirus						
Gastroenteritis	190.000	110	1217	1327	176	287
Campylobacter spp.						
Gastroenteritis	59,400	390	420	810	148	538
Guillain-Barré syndrome	60	35	150	185	150	185
Reactive arthritis	864	-	40	40	40	40
Inflammatory Bowel Disease	22	-	102	102	102	102
Total				1137		865
Salmonella spp.						
Gastroenteritis	35,400	440	255	697	77	517
Reactive arthritis	460	-	17	17	17	17
Inflammatory Bowel Disease	7	-	33	33	33	33
Total				747		567
STEC O157						
Gastroenteritis	1,300	6	13	19	7	13
Hemolytic Uremic Syndrome	20	76	25	101	25	101
Total				120		114

Discussion

The results showed that for three health states associated with foodborne disease less than 50% of the panel members were willing to trade-off any time. Therefore these health

states did not meet the relevance criterion proposed by Bonsel et al. [11]. Application of the preference-based relevance criterion reduced the burden of disease estimates of all five foodborne pathogens, varying from 94% (norovirus) to 5% (STEC O157). The ranking of the foodborne pathogens changed considerably when the relevance criterion was applied.

The burden of disease would decrease even more when disability weights based on median rather than mean TTO values were used to calculate YLD. The benefit of using median TTO values is that the majority rules principle is applied to all health states and not only minimal disease.

In our study the three health states that did not meet the preference-based relevance criterion of median TTO > 0 all had length-of-illness of one week or less and this short duration may have resulted in a majority of participants not willing to trade-off any time. Nonetheless, the results of two similar health state valuation studies showed that several health states that lasted over one week did not meet the preference-based relevance criterion either [16, 19]. Of the twelve health states that had a median TTO of 0, eight had a length-of-illness of over one week, varying from two weeks (genital candidiasis female) through 52 weeks (onychomicosis), whereas health state with a length-of-illness of one week or less on the other hand did meet the criterion [16]. This indicates that both duration and symptom severity matter for the panel of judges and that health state valuation studies are necessary to determine which health state meet the TTO median > 0 criterion.

In The Netherlands, intestinal infectious disease rank among the least burdening conditions and the results of this study might give the impression that application of the relevance criterion will substantially diminish it even further compared to other diseases [7]. However, it should be noted that of the many pathogens causing foodborne disease the current study addressed five. Not only is there a difference in disease caused by the foodborne pathogens, the severity and duration of the disease is also highly dependent of the condition of the patient. So for each cause of disease, foodborne or other, it should be carefully considered which health states are caused by the pathogen and whether all patients meet the health state descriptions.

Previously, relevance criteria based on anatomy and health care use have been suggested [20]. A problem with anatomical criteria is, however, that they cannot be applied to every condition and that for each group of diseases a specific anatomical criterion has to be formulated by experts. On the other hand, criteria based on health care use may be biased by differences in access to health care, resulting in incomparable disease burden [21]. Moreover, both anatomical and health care use criteria might be stringent and not allow relevant mild diseases to be included in burden of disease calculations. This might result in incomplete burden of disease estimates, an issue that is especially of importance for conditions characterized by heterogeneous levels of severity like foodborne disease and injury. This is underscored by the results of a recent study that showed a 36% increase in burden of disease if relevant mild injuries were included in burden on injury calculations [19].

The aforementioned problems are overcome by the preference-based relevance criterion. This method to empirically derive a relevance criterion is simple and transparent, and the resulting relevance criterion can be applied to each condition. Moreover, the relevance criterion is based on individual preferences derived from a population panel, which concurs with the societal perspective of the burden of disease concept [22]. Furthermore, as shown by the results of this study, the preference-based relevance criterion is sensitive for relevant mild disease.

A limitation of the preference-based criterion is that for any new health state preferences have to be obtained in order to derive the preference-based criterion. On the other hand, the results of the current study point out that preferences obtained from new panel studies collate with preferences derived previously with the same protocol.

Furthermore, it should be noted that the TTO preferences from this study are derived with the annual profile approach in stead of the standard QALY/DALY approach used in the original DALY-approach of the Global Burden of Disease study [3]. Unlike the annual profile approach, the standard QALY/DALY approach assumes independence between duration and disability and requires that the health state remain fixed over time. This means that in order to assess preference weights health states with an acute onset, episodic diseases like epilepsy, and health states characterized by complex and heterogeneous recovery patterns, have to be separated into numerous parts.

To alleviate this inability to assess preference weights for health states with complex patterns of duration and severity, the annual profile approach was developed [4]. The most important feature of the annual profile approach is that the course of the health state is described over one year time. The results of current and previous studies demonstrated that the annual profile approach yields valid and reliable disability weights for stable health states as well as health states that vary widely over time [13, 16, 19]. In absence of an algorithm to transform scores into utility values, the annual profile requires new panel data when new disease states have to be values. Yet, this does not imply that an infinite amount of states needs to be valued. The actual number required depends on the observed variability of severity-duration combinations. In most diseases less than ten states will suffice to cover the known heterogeneity of the disease. A second criticism of the annual profile approach was that it would overvalue diseases with a rapid course [23]. According to this criticism, application of these disability weights might overestimate the burden of disease. In their turn, Essink-Bot & Bonsel pointed out that this overvaluation of health states are a result of discrimination between low severity conditions rather than time presentation [4].

This alleged lack of discrimination between low severity conditions is not endorsed by the results of the current study, which show that the population panel in The Netherlands assigned values to health states that include the low severity conditions increase by level of severity of the health state. Additionally, they appeared to be capable to discriminate minimal from relevant disease and the preference-based relevance criterion appeared to be stable over time. Moreover, in Janssen et al. (2008) showed that the criterion is similar for lay people, medical advisors, as well as general practitioners [16].

A major limitation of the annual profile approach is that for any possible duration of a health state separate disability weights have to be elicited.
Conversely to the agreement between panels with different perspectives, it is not yet clear whether the preference-based relevance criterion is similar for different countries. Diseases that did not meet the relevance criterion in this study could be regarded as relevant by people who do not have good access to good sanitary facilities and health care. Previous studies did find that ranking of health states is similar across countries, yet the assigned values differed significantly [4, 24, 25]. Since the preference-based relevance criterion is based on values, rather than ranking, differences concerning the criterion are expected.
Secondly, it remains to be investigated whether the approach used to derive the preferences affects the preference-based relevance criterion. In the current study, the annual profile approach was used. The alternative standard QALY/DALY approach, unlike the annual profile approach, presents diseases with a so-called period profile. Using the standard QALY/DALY approach to obtain preferences might cause a shift in the derived preferences and consequently affect the preference-based relevance criterion.

We conclude that individual preferences derived with the TTO method from a population panel can be used to empirically derive a relevance criterion for burden of disease estimates, and that application of this preference-based relevance criterion results in considerable changes in ranking of foodborne pathogens.

Acknowledgements

We are grateful to Claudia Stein of the World Health Organization, Department of Food Safety, Zoonoses and Foodborne Diseases, for help and advice regarding the drafting of the manuscript.

References

1. Worldbank, World Development Report 1993: Investing in Health. 1993, New York: Oxford University Press.
2. Murray, C.J.L., A.D. Lopez, and C.D. Mathers, Summary measures of population health: concepts, ethics, measurement and applications. 2002, Geneva: World Health Organization.
3. Murray, C.J., J.A. Salomon, and C. Mathers, A critical examination of summary measures of population health. Bull World Health Organ, 2000. 78(8): p. 981-94.
4. Essink-Bot, M.L. and G.J. Bonsel, How to derive disability weights?, in Summary measures of population health: concepts, ethics, measurement and applications., C.J.L. Murray, A.D. Lopez, and J.A. Salomon, Editors. 2002, World Health Organization: Geneva.
5. Jankovic, S., et al., The burden of disease and injury in Serbia. Eur J Public Health, 2007. 17(1): p. 80-5.
6. Polinder, S., et al., Assessing the burden of injury in six European countries. Bull World Health Organ, 2007. 85(1): p. 27-34.

7. Melse, J.M., et al., A national burden of disease calculation: Dutch disability-adjusted life-years. Dutch Burden of Disease Group. Am J Public Health, 2000. 90(8): p. 1241-7.
8. Havelaar, A.H., et al., Health burden in the Netherlands due to infection with thermophilic Campylobacter spp. Epidemiol Infect, 2000. 125(3): p. 505-22.
9. Drummond, M., et al., Methods for the Economic Evaluation of Health Care Programmes. 1997, Oxford, U.K.: Oxford Medical Publications.
10. Brazier, J., M. Deverill, and C. Green, A review of the use of health status measures in economic evaluation. J Health Serv Res Policy, 1999. 4(3): p. 174-84.
11. Bonsel, G.J., M.F. Janssen, and E. Birnie, Mild diseases and ailments study (MiDAS). Ziektelastbepaling van 27 vermoedelijk 'lichte' ziekten ten behoeve van rationele beperking van het geneesmiddelenpakket. 2003, Academic Medical Center, University of Amsterdam: Amsterdam.
12. Krabbe, P.F., et al., The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. J Clin Epidemiol, 1999. 52(4): p. 293-301.
13. Stouthard, M.E., M.L. Essink-Bot, and G.J. Bonsel, Disability weights for diseases. A modified protocol and results for a Western European region. European Journal of Public Health, 2000. 10: p. 24-30.
14. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group Health Policy, 1990. 16(3): p. 199-208.
15. Janssen, M.F., E. Birnie, and G.J. Bonsel, Comparing the standard EQ-5D three level with a five level version. Value in Health, 2008. 11(2): p. 275-84.
16. Janssen, M.F., E. Birnie, and G.J. Bonsel, Feasibility and reliability of the annual profile model for deriving QALYs for short term health conditions. Medical Decision Making, 2008. 28(4): p. 500-10.
17. Torrance, G.W., W.H. Thomas, and D.L. Sackett, A utility maximization model for evaluation of health care programs. Health Serv Res, 1972. 7(2): p. 118-33.
18. Kemmeren, J.M., et al., Priority setting of foodborne pathogens. Disease burden and costs of selected enteric pathogens. 2006, RIVM: Bilthoven.
19. Haagsma, J.A., et al., Novel empirical disability weights to assess the burden of non-fatal injury. Injury Prevention., 2008. 14(1): p. 5-10.
20. Polinder, S., et al., International variation in clinical injury incidence: exploring the performance of indicators based on health care, anatomical and outcome criteria. Accident Analysis and Prevention, 2008. 40: p. 182-191.
21. de Wit, M.A., et al., Gastroenteritis in sentinel general practices,The Netherlands. Emerg Infect Dis, 2001. 7(1): p. 82-91.
22. Field, M.J. and M.R. Gold, Summarising population health: Directions for the development and application of population health metrics. . 1998, Washington D.C.: Institute of Medicine: National Academy Press.
23. Vos, T. The case against annual profiles for the valuation of disability weights., in Summary measures of population health: concepts, ethics, measurement and applications., C.J.L. Murray, A.D. Lopez, and J.A. Salomon, Editors. 2002, World Health Organization: Geneva.
24. Ustun TB, et al.: Multiple-informant ranking of the disabling effects of different health conditions in 14 countries. WHO/NIH Joint Project CAR Study Group. Lancet 1999. 354(9173): p. 111-5.
25. Schwarzinger, M., et al., Cross-national agreement on disability weights: the European Disability Weights Project. Popul Health Metr, 2003. 1(1): p. 9.

12

Discussion

This thesis addressed burden of disease measurement, in particular the methodology of the DALY. The DALY, which summarizes mortality and different disease end-points in a single figure, has revolutionized public health, because for the first time the public health effect of diseases and injuries could be compared in a systematic and quantitative way. Previously, cause specific mortality statistics and incidence data were used to indicate changes in the health status of a single country. With the DALY, health effects between different diseases, between subgroups of a population and even between countries can be compared. The Global Burden of Disease 1996 study used the DALY to estimate the total burden of disease worldwide as the sum of the burden of all separate diseases [1]. The results of this endeavour emphasized more than ever the inequalities that exist between countries and regions around the globe.

After its use in the groundbreaking Global Burden of Disease (GBD) studies, the DALY gained wide adherence, aiding national and international decision making processes regarding resource allocation and prevention [1, 2]. In the Netherlands, the DALY has been adopted as the standard outcome measure in public health and the primary outcome for priority setting in health care and prevention since its first application in the Dutch Public Health Status and Forecasts study in 1997 [3, 4].

However, apart from the obvious advantages, the DALY concept has been criticized concerning key aspects of the methodology [5-7]. This thesis addressed several of these controversial aspects with regards to two health (care) domains where these aspects are particularly problematic. Both health domains, injuries and infectious intestinal disease, are featured by heterogeneous health outcomes, including the extremes of the severity spectrum and duration, with all types of time-severity relations. Within these health domains alternative solutions for four controversial methodological aspects have been examined. When satisfactory solutions for these controversial methodological aspects have been found, the DALY can proceed its triumphal march in public health. However, if the alternative solutions are not proven satisfactory, the use of the DALY in these domains of public health, and possible other domains with similar methodological constraints, should be reconsidered.

The main questions of this thesis were:

- How can the assessment of health-related quality of life and the process to derive disability weights regarding acute onset disorders with a highly variable course over time be improved?
- Which of three existing comorbidity approaches in DALY calculations performs best?
- What is the impact of commonly ignored long-term sequelae of acute onset disorders on burden of disease estimates?
- How can a threshold of triviality be developed to separate cases of acute onset disorders which are relevant for public health policy from those which can validly be ignored regardless of their number (prevalence)?

Regarding these questions this thesis provided the following answers:

- The assessment of health-related quality of life and disability weight for acute onset disorders regarding acute onset disorders with a highly variable course over time can be improved by 1) using the annual profile approach, 2) including disease specific health state descriptions, especially in case of conditions that affect physical appearance, and 3) when a MAUI instrument is used, using a combination of MAUI instruments to cover all relevant health dimensions.
- The performance of available comorbidity approaches is reasonable and regarding the relative frequent occurrence of comorbidity in the general population, comorbidity adjustment should be applied to avoid wrongful policy recommendations regarding priority setting and prevention.
- A strong association between acute onset disorders and previously ignored long-term sequelae was found and inclusion of these sequelae increases the burden of disease of these disorders tremendously.
- In burden of disease estimates, cases of minimal disease in YLD estimates may be identified, measured and accounted for by using a relevance criterion. The threshold of triviality may be based on health care, anatomical and outcome criteria. The preference-based relevance criterion, a criterion that is based on outcome, distinguishes experienced minimal disease from relevant disease unambiguously, yet it also allows relevant mild disease to be included in burden of disease estimates in order to avoid incomplete estimates.

The four principal questions and our answers will be discussed in the following sections.

Disability weights and health-related quality of life

Research question 1: How can the assessment of health-related quality of life and the process to derive disability weights for acute onset disorders with a highly variable course over time be improved?

Main outcomes

This thesis showed that the annual profile approach, which avoids the breakdown of health outcome into time periods of constant health, allows the derivation of disability weights for acute onset disorders with a highly variable course over time. Disorders that have been ignored in previous burden of injury estimates. Alternative to the annual profile approach, patient-derived follow-up MAUI data may be used to derive otherwise lacking disability weights for acute onset disorders. However, comparing annual profile disability weights to MAUI disability weights for injury consequences revealed a possible overestimation of temporary consequences of minor injuries (i.e. injuries of low severity) when MAUI disability weights were used. The relative high values of the MAUI disability weights for mild injuries in combination with high incidence results in a large number of total YLDs, and this may lead to policy priority of these mild injuries above severe,

less frequently occurring injury consequences. These results indicate that the MAUI approach to derive disability weights is inappropriate for injury consequences and temporary consequences of low severity injuries in particular.

Comparison of the annual profile approach to alternative options

Apart from the underlying theoretical assumptions of constancy of health that do not fit, practical objections hamper the value of the MAUI approach to assess disability weights for injury consequences with complex and heterogeneous patterns over time. For injuries, the first follow-up moment at which the health status of injury patients is measured, e.g. 2.5 months post-injury in the Dutch follow-up study, is often inappropriate for this purpose [8]. At 2.5 months, many patients with temporary consequences have already recovered largely or even completely from their injury. This means that the actual health effects of temporary consequences are not picked up on. Measuring the health status of injury patients with a shorter follow-up time might solve this problem. The UK burden of injury study measured the health status of injury patients in four UK areas one month, four months and 12 months post-injury [9]. At one month follow-up, it is expected that a smaller proportion of patients has recovered compared to the 2.5 month assessment of the Dutch follow-up study. Nevertheless, the data from UK burden of injury study data meets constraints similar to those that are met the Dutch follow-up study. With both studies, the time intervals between the three fixed follow-up moments are relatively long, namely several months). When these data are used to derive MAUI disability weights, constant health is assumed for relatively long periods of time and this may lead to an overestimation of the actual impact of the injury consequences.
Apart from this, information on pre-injury data is needed to assess the true impact of injury on health-related quality of life. In this thesis UK population summary scores were used to calculate EQ-5D disability weights [10], because Dutch population summary scores are lacking. The benefit of using population summary scores as a baseline is that the disability weights can be adjusted for age and sex of the patient. New evidence suggests that EQ-5D population scores are significantly lower than pre-injury EQ-5D scores among injury patients [9]. This implies that using population scores for the calculation of EQ-5D disability weights results in an underestimation of the impact of the injury. Therefore, pre-injury health-related quality of life data should be used as a baseline for the EQ-5D disability weight calculations. However, pre-injury health-related quality of life data are often not available. The Dutch follow-up study, for instance, did not measure pre-injury health-related quality of life [8]. The UK burden of injury study and an Australian study have measured pre-injury data, though the validity of these pre-injury data are questionable [9, 11]. In these studies, pre-injury health-related quality of life was assessed retrospectively, i.e. after the injury patients experienced the shock of accidentally sustaining an injury and the impact of its consequences on daily life. This may have affected pre-injury health-related quality of life measurement.

Other drawbacks of the MAUI approach have also been found. This thesis demonstrated a large disparity between disease specific and generic health state valuations. This disparity was particularly apparent in case of minor disorders and disorders that affect physical appearance. This indicates that the generic MAUI description does not contain information on prominent symptoms that are reflected in the disease specific description, affecting health state valuation and eventually the resulting disability weight.
However, the amount of information loss is dependent on the attributes that are incorporated in the MAUI instrument. Among patients with long-term injury consequences (sequelae), a comparison of patient-reported HUI2, HUI3 and EQ-5D data showed that the HUI is more sensitive compared to the EQ-5D. The EQ-5D does not measure disability, where the HUI does. The HUI for instance is more sensitive for comorbid disease. The sensitivity of the instrument to certain key symptoms is therefore highly dependent on the instrument that was used to derive the health-related quality of life data. For instance, the EQ-5D data is not equipped with a cognitive attribute, whereas the HUI is [12, 13]. This means that, opposed to the HUI, in case of brain injury the possible decrease of cognitive abilities cannot be assessed with the EQ-5D. Apart from the attributes, the performance of the instrument is also dependent on the number of response categories, which is higher in case of HUI. This inevitably results information loss, reflected by the higher percentage of missed disability with EQ-5D compared to HUI[14].

Strengths and weaknesses of the annual profile approach

Are annual profile disability weights a solution? This thesis showed that the annual profile approach alleviates several constraints of the standard approach, diminishing the problems that are raised by the latter [15, 16]. Furthermore, information loss is less an issue with the annual profile approach, because it describes the health state using both disease specific and generic information [15].
A disadvantage of the annual profile approach is that the presented descriptions of the health state contain a large amount of information. This makes these descriptions conceptually difficult, especially in conjunction with the health state valuation task. Apart from the difficulty of the valuation task, the annual profile approach is also debated regarding its practical disadvantage of having to elicit separate weights for any possible severity state and duration. This does not imply that an infinite amount of states needs to be valued. The actual number that is required depends on the observed variability of severity and duration combinations. For most diseases less than ten health states will suffice to cover the known heterogeneity of the disease. However, in absence of an algorithm to transform scores into utility values, the annual profile requires new panel data when new disease states have to be valued.
This has proved to be a significant problem, since the linkage to injury epidemiological data has shown that the current set of available annual profile disability weights for injury

consequences does not suffice. To calculate YLD, injury incidence data has to be linked to disability information and for this linkage the injury incidence is coded into injury groupings, for instance by using the EUROCOST injury diagnosis classification scheme. EUROCOST identifies 39 injury groups that may then be regrouped into ten broader categories [8, 17]. This means that for each of these injury groups appropriate disability weights are needed. Linkage of injury cases in ED and hospital discharge records showed that the percentage of cases that can be coded with EQ-5D disability weights is approximately 90%. However, for the annual profile approach disability weights with approximately 65% of coded cases this percentage is much lower [18].

Application of the annual profile approach

Combining the results of these studies revealed that the neglect of acute temporary health outcome results in an underestimation of the burden of disease, yet applying unfit MAUI or period profile disability weights results in an overestimation. Both options result in highly inaccurate estimates of the burden of disease, restricting the use of these studies in priority setting of resources.
The annual profile approach proved to be a satisfactory solution for this problem. In practice, for injuries and infectious intestinal disease this meant that new annual profile disability weights had to be derived for acute onset disorders and that these disability weights had to be applied in burden of disease estimates.
Previously, in burden of injury studies acute temporary consequences of injury have been ignored because of lack of valid disability weights. Application of the novel set of disability weights in the Dutch burden of disease and injury study showed that inclusion of these health outcomes resulted in an increase of non-fatal burden of injury of 36%. This demonstrates that the application of the annual profile disability weights can improve the estimates of the non-fatal burden of acute onset disorders, such as injuries and intestinal infectious disease.

Recommendations for future research

- The current set of annual profile disability weights for injury consequences does not suffice, because for a group of injury consequences disability weights are lacking. To enhance linkage of epidemiological registry records and annual profile disability weights an additional set of annual profile disability weights for injury consequences should be derived.
- Annual profile disability weights for other health domains featured by heterogeneous health outcomes with all types of time-severity relations have to be derived order to improve burden of disease estimates in these health domains.
- When the MAUI approach is used to derive disability weights, it is recommended that they are used for stable, chronic health states only.
- To cover all relevant health dimensions, a combination of the HUI and EQ-5D instrument in studies on disability should be used. Data from both instrument may

than be used to derive disability weights with the MAUI approach and possibly for deriving annual profile disability weights as well [19].

- Furthermore, information on pre-disease/injury data should be gathered to assess the true impact of the health outcome on health-related quality of life.
- For future disability weights studies that incur panel elicitation, it is recommended to assess information load and its effect on processing and the health state valuation task.

Comorbidity Adjustment

Research question 2: Which of three existing comorbidity approaches in DALY calculations performs best?

Main outcomes

The existing three approaches to deal with the impact of comorbidity (i.e. the maximum limit, additive and multiplicative approach) all involve the disability weight. Follow-up data from injury patients provided an opportunity to test the impact of comorbid disease and short-term injury, because of the temporal features of the injury consequences. This thesis demonstrated that the three comorbidity approaches to account for comorbidity effects provided reasonable results.

Alternative solutions for comorbidity adjustment

In addition to further study of the performance of currently available comorbidity approaches, other adjustment approaches should be developed. One such approach is a domain specific approach. This approach differs from the calculation EQ-5D utility scores that was examined, because it starts from the domain specific impact of a condition without comorbidity and compares this impact with the estimated domain impact of the comorbidity only. This approach can accommodate co-existing diseases which share affected domains, or the presence of two or more comorbid diseases. It does, however, require detailed descriptive data from a large number of patients.
A second alternative method to adjust for comorbidity is to adjust life expectancy for the reduced life span due to other diseases when a certain disease is associated with (severe) comorbidity.

Advantages and disadvantages of comorbidity adjustment

Previously, the existing comorbidity approaches have been tested regarding their effect on DALY estimates [20, 21]. In this thesis, comorbidity adjustment was examined using actual patient data to test each of the adjustment methods.

A disadvantage of this patient data was that the generic 2.5 months EQ-5D data of the Dutch follow-up study was used to examine the comorbidity adjustment methods [8]. As mentioned above, at 2.5 months many patients with injuries have already recovered largely or completely from their injury.
Performing a similar study on patient data with shorter follow-up time and consequently less patients that recovered from their injury might yield different results regarding the performance of the comorbidity adjustment approaches.
A second disadvantage of the procedure used in this thesis is that the sensitivity of the EQ-5D, the MAUI instrument that was used to assess the effect of injury and chronic disease, proved to be low compared to other MAUI instrument [14]. The HUI is more sensitive for comorbid disease compared to the EQ-5D and that the EQ-5D does not measure disability, where the HUI does. To strengthen the insight concerning the current approaches to adjust for comorbidity, we therefore recommend more research on the effects of comorbidity on health-related quality of life by examining health-related quality of life measured with the HUI.
More information is also needed on whether the approaches hold if multiple chronic diseases coexists.

Application of comorbidity adjustment

Comorbidity is highly prevalent in the population. Among injury patients approximately one third of patients have one or more diseases prior to the injury [22-26]. In current burden of disease studies, estimates have not been adjusted for comorbidity. This neglect has implications for priority setting and prevention, especially if these diseases complicate the prognosis of secondary diseases. In case a disorder is secondary more often than primary, their significance in burden of disease studies and the benefit of eradication is underestimated.
This thesis showed that the performance of available comorbidity adjustments methods is reasonable and that they may be applied in burden of disease studies. Moreover, regarding the relative frequent occurrence of comorbidity in the general population, comorbidity adjustment should be applied to avoid wrongful policy recommendations regarding priority setting and prevention.

Recommendations for future research

- Regarding comorbidity adjustment approach, it is recommended to study the development of other comorbidity approaches, such as the domain specific approach. The domain specific approach can accommodate co-existing diseases which share affected domains.
- It is recommended to study the performance of the comorbidity approaches with patient data assessed at a shorter follow-up time to eliminate the effect of recovery from injury.

- To strengthen the insight in current approaches to adjust for comorbidity, more research is needed on the effects of comorbidity on health-related quality of life measured with the HUI and/or other MAUI instruments.
- More information is needed on the performance of the comorbidity approaches in case of multiple chronic diseases coexist.

Long-term sequelae

Research question 3: What is the impact of commonly ignored long-term sequelae of acute onset disorders on burden of disease estimates?

Main outcomes

This thesis demonstrated a strong association between post-infectious irritable bowel syndrome and infectious intestinal disease. Post-infectious irritable bowel syndrome manifests itself in 9% of patients with bacterial infectious intestinal disease and accounts for half of the total disease burden for the selected pathogens.
Regarding the health domain injury, PTSD was shown to be relatively prevalent among injury patients. Shorter than 3 months after the injury a PTSD prevalence rate of 24% was found, tapering down to 4% at 24 months after the injury. Including PTSD increased the non-fatal burden of disease of unintentional injuries with 54%.
These findings point out that neglecting sequelae in burden of injury estimations leads to a major underestimation and that it is quintessential to define all possible health outcomes of a disease or injury, including remote and psychological consequences, to establish the causal relation and include them in the burden of disease estimates.

Alternatives to in- and exclusion criteria of health outcomes

Choices regarding the in- and exclusion of long-term sequelae may be based consciously on preliminary estimations of the relative impact of all possible health outcomes on the total disease burden by excluding outcomes that contribute little to the burden of disease because they are extremely rare, or because their impact on health-related quality of life is low [27]. However, more often health outcomes are excluded unconsciously because researchers are unaware of certain sequelae and/or because data on sequelae are lacking. The latter is particularly evident in case a sequel manifest itself relatively long after the patient recovered from the acute disorder.
For burden of disease studies that aid priority-setting in health care and prevention, the burden of disease estimate should comprise a complete assessment of the health effects due to a certain cause. To facilitate researchers in the area of burden disease studies, the GBD drew up a list of possible sequelae [28]. By doing this, the WHO underlined the

importance of complete burden of disease estimates. Nonetheless, a close investigation learns that marked sequelae are omitted, whereas sequelae that are less likely to occur are included, confirming the sense of arbitrariness of the underlying decision rules concerning the in- and exclusion criteria of sequelae.
This sense of arbitrariness became particularly apparent regarding the health domain of injury, where the inclusion criteria of sequelae seemed to be constricted to physical consequences, whereas psychopathological consequences, such as PTSD, travel anxiety and other maladaptive syndromes - undisputed consequences of high impact once they occur - are not taken into account.
The use of a framework to map where possible effects of a disease or injury can be expected and a clearly defined procedure to assess sequelae, such as the construction of outcome trees, may help to identify otherwise ignored long-term sequelae.

Strengths and weaknesses of the investigation of long-term sequelae

The examination of long-term sequelae in general and in this thesis in particular is complicated, because conspicuous difficulties may hamper the evidential value of the sequelae. Difficulties such as time delay between acute disorder and onset of the sequelae, and the fact that sequelae may be relatively common in the general population.
Weaknesses that are specific for PI-IBS concern the epidemiology and course of PI-IBS. Follow-up studies should provide more information on the risk of IBS after bacterial and viral infectious intestinal disease and duration of the symptoms.
In the field of injuries conspicuous difficulties hamper the evidential value of the injury as actual cause of the psychological disorder. These difficulties are caused by the time delay between injury and onset of symptoms, as well as by the fact that many psychological disorders have a relatively high prevalence in the general population. Both difficulties were not addressed in the current thesis. However, new evidence revealed that approximately half of the injury patients diagnosed with PTSD experienced symptoms before the injury [29]. This highlights the importance of these aspects. Long-term case-control studies of injury patients may help gain more insight in the relation between injury and psychological disorders. This insight is needed to unravel the intricate combinations of factors that cause PTSD and aid in the attribution of its disease burden. If this aspect is not addressed, the disease burden of PTSD cases may be attributed to multiple causes and consequently double counted in burden of disease studies. Double counting is an issue in burden of disease studies that has to be reckoned with. To adjust for double counting counterfactual impact analysis may be used. With this analysis the observed burden of disease is attributed to risk factors. The use of counterfactual impact analysis, however, strongly depends on a solution of the comorbidity assignment problem.

Taking into account long-term sequelae: application in burden of disease studies

Calculations of the burden of injury have important restrictions, since they often focus on immediate physical outcomes only. However, diseases and injuries may be associated with

a range of outcomes including psychopathological, which fit the definition of a sequelae. To avoid wrongful policy recommendations regarding priority setting and prevention, all sequelae of a certain disease or injury should be included. Regarding PTSD, this means that the burden of PTSD that is caused by injury should be attributed to the injury category and not to a general category of anxiety disorders, which is currently the practice [1, 3, 30].

Furthermore, for some health domains the definition of sequelae needs to be revised, since it is often restricted to acute and/or physical consequences. In a recent paper, Lyons et al. (2009) presented a framework that comprises the injury burden across individual, family and societal domains, emphasizing the importance of broadening the definition of injury consequences [18]. Furthermore, this framework may be used to map where possible effects of an injury can be expected and it may be applicable to health domains other than injury.

Recommendations for future research

- For future burden of disease studies it is recommended to define all possible health outcomes of a certain injury or disease. These possible health outcomes should cover immediate as well as remote health outcomes, and physical as well as psychological health outcomes. This implies that for all health domains, the definition of sequelae should be broadened to physical as well as psychological health outcomes.
- A framework, such as the LOAD framework [18], may be used to map where possible effects of a disease or injury can be expected. Furthermore, a clearly defined procedure to assess sequelae, such as the construction of outcome trees, may help to identify otherwise ignored long-term sequelae.
- For health domains other than injury and infectious intestinal disease, long-term follow-up studies should be conducted as well to gain more insight in the development of long-term sequelae among patients with acute onset disorders.
- To adjust for background illness in the general population and to overcome the difficulties in the attribution of long-term sequelae to the disease or injury, case-control studies should be performed.
- To avoid double counting in burden of disease studies, decision rules regarding the attribution of sequelae should be developed, especially in case sequelae may be attributed to multiple causes.
- To enhance the burden of disease estimates of infectious intestinal disease in particular, follow-up studies are recommended to provide more information on the risk of developing the risk of developing IBS after viral infectious intestinal disease and the course of PI-IBS after viral and bacterial infectious intestinal disease.
- To enhance the burden of injury in particular, follow-up studies regarding injury patients are recommended to gain insight in predictors and course of PTSD, as well as the association between injury and psychological consequences other than PTSD, such as depression and travel anxiety.

Minimal disease

Research question 4: How can a threshold of triviality be developed to separate cases of minimal disease which are relevant for public health policy from those which can validly be ignored regardless their number (prevalence)?

Main outcomes

This thesis explored the performance of indicators based on health care, anatomical and outcome criteria to identifying cases that are relevant from a public health perspective.
An example of an indicator based on outcome criteria is the preference-based relevance criteria. The preference-based relevance criterion is met if the median TTO value is greater than 0, i.e. if at least 50% of a panel of judges is willing to trade-off time in order to be restored to full health. If not, the health state is regarded as not relevant and excluded from the burden of disease calculation.
This thesis demonstrated that the preference-based relevance criterion distinguishes experienced minimal disease from relevant disease unambiguously, yet it also allows relevant mild disease to be included in burden of disease estimates in order to avoid incomplete estimates.

Comparison to other relevance criteria

An alternative solution to distinguish between minimal and relevant disease may be to apply a relevance criterion that concerns health care use. In many burden of disease and injury studies, health care use (i.e. hospital admission, emergency department treatment or General Practitioner consultation) is used as a severity cut-off point [18]. However, the very nature of these health care facilities implies that not all cases of injuries are covered, due to self-selection and, to a small degree, selection at the hospital department. Moreover, differences in health care seeking behaviour between countries prevent the use of General Practitioner consultation as a threshold of relevant disease. Comparison of two studies regarding health care use in England and The Netherlands showed that patients suffering from infectious intestinal disease from England are three times more likely to consult a General Practitioner [31-35]. A recent study that reconstructed the surveillance pyramid of infectious intestinal disease in a large number of European countries revealed that the proportion of patients with gastroenteritis that consult the General Practitioner varies widely between countries [36].
Using a severity threshold based on an anatomical criterion may meet constraints as well. In a previous burden of injury study only injury patients with major trauma were included [37]. In this study, major trauma was defined by an injury severity score of 15 or higher, which is an internationally used definition of major trauma [38]. The application of this anatomical criterion allows researchers to focus on important injuries in terms of consequences. However, many injury consequences that are relevant according to injury

patients and population panels and important in terms of incidence, such as single fractures, are excluded when this criterion is applied. The preference-based relevance criterion, on the other hand, includes injuries that are regarded important by patients and population, yet highly incident irrelevant injuries, such as superficial injury and corneal abrasion are not considered.

Strengths and weaknesses of the preference-based relevance criterion

A great advantage of the preference-based relevance criterion is that it allows relevant mild disease to be included in burden of disease estimates in order to avoid incomplete estimates of the burden of disease. This is an issue that is especially of importance for conditions characterized by heterogeneous levels of severity like infectious intestinal disease and injury. This thesis also showed that the relevance criterion can be applied in several domains in health care. Moreover, the relevance criterion is based on individual preferences derived from a population panel, which concurs with the societal perspective of the burden of disease concept.

Nevertheless, cross-cultural differences may also exist in perceiving disease and as a result may effect the preference-based relevance criterion. Previous studies have shown that similar health states are valued differently by population panels of different countries. Health state valuations of the EQ-5D, for instance, have resulted in different value sets, even though a similar methodology was used to elicit the health state valuations [39]. Disease specific health state valuations showed that ranking of health states is similar across countries, yet the assigned values differed significantly [6, 40, 41]. Since the preference-based relevance criterion is based on values rather than ranking, differences concerning the criterion are expected. For instance, diseases, such as mild gastroenteritis, that did not meet the relevance criterion in this study could be regarded as relevant by people who do not have good access to good sanitary facilities and health care.

Application of the preference-based relevance criterion

The preference-based relevance criterion proved to be a satisfactory solution for the problem of identifying, measuring and accounting for cases of acute onset disorders in burden of disease estimates that are relevant for health policy. In The Netherlands, the preference-based relevance criterion has been applied in burden of disease studies of both injuries and infectious intestinal disease. For injuries, the application of the preference-based relevance criterion meant decrease of 3%. For infectious intestinal disease (pathogens norovirus, rotavirus, Salmonella, Campylobacter and STEC O157) the burden of disease decreased by 312%. Regarding the latter, it should be noted that post-infectious irritable bowel syndrome was not included in the burden of disease estimates.

For health domains other than injury and infectious intestinal disease that are featured by heterogeneous health outcomes and that include the extremes of the severity spectrum, the preference-based relevance criterion has yet to be derived.

In practice, this means that new health state valuations for health outcomes associated with these domains have to be elicited. The results of these health state valuations are then used to assess the preference-based relevance criterion in these health domains.

Recommendation for future research

- The health outcomes that we identified as irrelevant, such as mild gastroenteritis and superficial injury, are based on health state valuations of a Dutch laymen panel. It is recommended to investigate which health outcomes do not meet the preference-based relevance criterion for countries other than The Netherlands.
- Moreover, the preference-based relevance criterion should be assessed for health domains other than injuries and infectious intestinal disease that are featured by heterogeneous health outcomes as well.

Conclusions

Murray et al. have delivered a great achievement by developing the DALY concept [1]. From its launch in 1993, the DALY has revolutionized public health. Moreover, the majority of the methodology conceptualized by Murray et al. is still valid. However, in two instances the methodology of the DALY is inadequate. First, in case disorders are characterized by heterogeneous health outcomes which include the extremes of the severity spectrum and with all types of time-severity relations. Second, in case two or more disorders occur simultaneously in a person.
This thesis demonstrated satisfactory alternative solutions to controversial issues of the DALY concept. As a result, the DALY concept may be used for burden of disease studies in health domains where these controversial issues emerge in particular. However, the application of alternative solutions implies that the standard DALY approach should be abandoned. This standard approach is appropriate for chronic health outcomes, but unsuitable for intricate time-severity relations because of its underlying assumptions. Nonetheless, the standard DALY approach is often used for health outcomes characterized by complex time-severity relations as well. Application of alternative solutions proposed in this thesis may improve burden of disease estimates considerably.

References

1. Murray, C.J.L. and A.D. Lopez, The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. 1996, Cambridge: Harvard University Press.
2. Murray, C.J. and A.K. Acharya, Understanding DALYs (disability-adjusted life years). J Health Econ, 1997. 16(6): p. 703-30.

3. Melse, J.M., et al., A national burden of disease calculation: Dutch disability-adjusted life-years. Dutch Burden of Disease Group. Am J Public Health, 2000. 90(8): p. 1241-7.
4. Ruwaard D. and K.P.G. N., Volksgezondheid Toekomst Verkenning 1997. De som der delen. 1997, Utrecht: Elsevier/De Tijdstroom.
5. Anand, S. and K. Hanson, Disability-adjusted life years: a critical review. J Health Econ, 1997. 16(6): p. 685-702.
6. Essink-Bot, M.L. and G.J. Bonsel, How to derive disability weigths?, in Summary measures of population health: concepts, ethics, measurement and applications., C.J.L. Murray, A.D. Lopez, and J.A. Salomon, Editors. 2002, World Health Organization: Geneva.
7. Vos, T. and C.D. Mathers, The burden of mental disorders: a comparison of methods between the Australian burden of disease studies and the Global Burden of Disease study. Bull World Health Organ, 2000. 78(4): p. 427-38.
8. Polinder, S., et al., Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. J Trauma, 2007. 62(1): p. 133-41.
9. Lyons, R.A., et al., The UK burden of injury study - a protocol. [National Research Register number: M0044160889]. BMC Public Health, 2007. 7: p. 317.
10. Kind, P., G. Hardman, and S. Macran, UK population norms for EQ-5D. 1999, Centre of Health Economics, University of York: York.
11. Watson, W.L., J. Ozanne-Smith, and J. Richardson, Retrospective baseline measurement of self-reported health status and health-related quality of life versus population norms in the evaluation of post-injury losses. Inj Prev, 2007. 13(1): p. 45-50.
12. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group Health Policy, 1990. 16(3): p. 199-208.
13. Feeny, D., et al., Multi-attribute health status classification systems. Health Utilities Index. Pharmacoeconomics, 1995. 7(6): p. 490-502.
14. Bas Janssen, M.F., E. Birnie, and G.J. Bonsel, Evaluating the discriminatory power of EQ-5D, HUI2 and HUI3 in a US general population survey using Shannon's indices. Qual Life Res, 2007. 16(5): p. 895-904.
15. Janssen, M.F., E. Birnie, and G. Bonsel, Feasibility and reliability of the annual profile method for deriving QALYs for short-term health conditions. Med Decis Making, 2008. 28(4): p. 500-10.
16. Stouthard, M.E., M.L. Essink-Bot, and G.J. Bonsel, Disability weights for diseases. A modified protocol and results for a Western European region. European Journal of Public Health, 2000. 10: p. 24-30.
17. Polinder, S., et al., Cost estimation of injury-related hospital admissions in 10 European countries. J Trauma, 2005. 59(6): p. 1283-90; discussion 1290-1.
18. Lyons, R.A., et al., The Injury List Of All deficits (LOAD) Framework - conceptualising the full range of deficits and adverse outcomes following injury and violence. Int J Inj Contr Saf Promot, 2009: p. 1-15.
19. Van Beeck, E.F., et al., Guidelines for the conduction of follow-up studies measuring injury-related disability. J Trauma, 2007. 62(2): p. 534-50.
20. van Baal, P.H., et al., Disability weights for comorbidity and their influence on health-adjusted life expectancy. Popul Health Metr, 2006. 4: p. 1.
21. Mathers, C.D., K.M. Iburg, and S. Begg, Adjusting for dependent comorbidity in the calculation of healthy life expectancy. Popul Health Metr, 2006. 4: p. 4.
22. Aggarwal, R., et al., Psychometric Properties of the EuroQol-5D and Short Form-6D in Patients with Systemic Lupus Erythematosus. J Rheumatol, 2009.
23. Barton, G.R., et al., Comparing Utility Scores Before and After Hearing-Aid Provision : Results According to the EQ-5D, HUI3 and SF-6D. Appl Health Econ Health Policy, 2004. 3(2): p. 103-5.
24. Hoeks, S.E., et al., Clinical validity of a disease-specific health status questionnaire: the peripheral artery questionnaire. J Vasc Surg, 2009. 49(2): p. 371-7.

25. McCrone, P., et al., A Comparison of SF-6D and EQ-5D Utility Scores in a Study of Patients with Schizophrenia. J Ment Health Policy Econ, 2009. 12(1): p. 27-31.
26. Selai, C.E., K. Elstner, and M.R. Trimble, Quality of life pre and post epilepsy surgery. Epilepsy Res, 2000. 38(1): p. 67-74.
27. Havelaar, A.H., Methodological choices for calculating the disease burden and cost-of-illness of foodborne zoonoses in European countries. 2007, Network for the Prevention and Control of Zoonoses.
28. The global burden of diseases, injuries, and risk factors study: operations manual. . 2009 [cited.
29. Bryant, R.A., et al., The psychiatric sequelae of traumatic injury. Am J Psychiatry. 167(3): p. 312-20.
30. van der Lucht, F. and J.J. Polder, Van gezond naar beter. Kernrapport van de Volksgezondheid Toekomst Verkenning 2010. 2010, Rijksinstituut voor Volksgezondheid en Milieu: Bilthoven.
31. de Wit, M.A., et al. Gastroenteritis in sentinel practices in the Netherlands. in 4th World Congress Foodborne Infections and Intoxications. 1998.
32. de Wit, M.A., et al., Etiology of gastroenteritis in sentinel general practices in the netherlands. Clin Infect Dis, 2001. 33(3): p. 280-8.
33. de Wit, M.A., et al., Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol, 2001. 154(7): p. 666-74.
34. Tompkins, D.S., et al., A study of infectious intestinal disease in England: microbiological findings in cases and controls. Commun Dis Public Health, 1999. 2(2): p. 108-13.
35. Wheeler, J.G., et al., Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive. Bmj, 1999. 318(7190): p. 1046-50.
36. Havelaar, A.H., WP 23: Prioritising foodborne and zoonotic hazards at EU level. Deliverable D23.06: Summary of available data. 2009, Network for the Prevention and Control of Zoonoses.
37. Holtslag, H.R., et al., Individual and population burdens of major trauma in the Netherlands. Bull World Health Organ, 2008. 86(2): p. 111-7.
38. Gebhard, F. and M. Huber-Lang, Polytrauma--pathophysiology and management principles. Langenbecks Arch Surg, 2008. 393(6): p. 825-31.
39. Szende, A., M. Oppe, and N.E. Devlin, EQ-5D value sets: Inventory, Comparative Review and User Guide EuroQol Group Monographs. Vol. 2. 2007, Dordrecht: Springer.
40. Fewtrell, L.J., et al., Estimating the global burden of disease of mild mental retardation and cardiovascular diseases from environmental lead exposure. Environ Res, 2004. 94(2): p. 120-33.
41. Schwarzinger, M., et al., Cross-national agreement on disability weights: the European Disability Weights Project. Popul Health Metr, 2003. 1(1): p. 9.

Summary

The population's health faces an array of diseases and injuries. Limited resources compel policy-makers everywhere to focus on threats that are regarded most relevant in terms of public health. The World Health Organization (WHO) and Worldbank developed an innovative concept, which expresses the burden of disease in Disability Adjusted Life Years (DALY). The DALY provides knowledge on the size of health problems and the potential benefit of proposed measures set against similar and comparable data of other problems.
However, apart from the obvious advantages, the DALY concept has been criticized concerning key aspects of the methodology. This thesis addressed four of these controversial aspects with regards to two health (care) domains where these aspects are particularly problematic. Both health domains, injuries and infectious intestinal disease, are featured by heterogeneous health outcomes, including the extremes of the severity spectrum and duration, with all types of time-severity relations.

1. Disability weights and health-related quality of life

The first aspect concerns the process to derive disability weights. The disability weight is a value assigned to living with disability and is anchored between 0 and 1. Problematic, however, is that in the original DALY concept independence between duration and disability is assumed, which requires health state to remains constant over time. This assumption is untenable for many diseases and injury consequences. Moreover, disability weights for highly frequent temporary consequences of injury were not lacking.
In this thesis, a set of 43 disability weights for both permanent and temporary consequences of injury were derived using annual profile approach (APA). The APA describes the course of the condition over one year time, allowing assessment of disability weights for health states characterized by an acute onset and complex patterns of recovery. A population panel (n = 143) provided the values. Application of these new disability weights showed a 36% increase in YLD due to unintentional injury.
Alternative to the annual profile approach, patient-derived follow-up data derived with multi-attribute utility instruments (MAUI) might be used to derive otherwise lacking disability weights for acute onset disorders. A MAUI classify health states with generic attributes, which is then converted into weight that fits within the 0-1 range. In this thesis disability weights were assessed using data derived with the MAUI EQ-5D from injury patients 2½, 5, 9 and 24 months after attending the Emergency Department. Comparing these MAUI disability weights to the APA disability weights revealed a possible major overestimation of temporary consequences of minor injuries (i.e. injuries of low severity) when MAUI disability weights were used. This is possibly a result of the underlying theoretical assumption of constancy of health that do not fit. The relative high values of the MAUI disability weights for mild injuries in combination with high incidence result in a large number of total DALYs, and may lead to policy priority of these mild injuries

above severe, less frequently occurring injury consequences. These results indicate that the MAUI approach to derive disability weights is inappropriate for injury consequences and temporary consequences of low severity injuries in particular.
A comparison of generic EQ-5D and disease specific health state valuation by a population panel revealed a large disparity between disease specific and generic health state valuations. This disparity was particularly apparent in case of minor disorders and disorders that affect physical appearance. This indicates that the generic MAUI description does not contain information on prominent symptoms that are reflected in the disease specific description, affecting health state valuation and eventually the resulting disability weight.
The amount of information loss is, however, dependent on the attributes that are incorporated in the MAUI that is used. Among patients with long-term injury consequences (sequelae), a comparison of patient-reported data derived with the MAUIs EQ-5D and the Health Utilities Index (HUI) showed that the HUI is more sensitive compared to the EQ-5D and that the EQ-5D does not measure disability, where the HUI does. The HUI for instance is more sensitive for comorbid disease. The sensitivity of the instrument to certain key symptoms is therefore highly dependent on the instrument used to derive the patient data.

2. Comorbidity adjustment

Comorbidity is highly prevalent in the population. In current burden of disease studies, estimates have not been adjusted for comorbidity. This neglect may lead to wrongful policy recommendations regarding priority setting and prevention. EQ-5D follow-up data from injury patients provided an opportunity to test the impact of comorbid disease and short-term injury, because of the temporal features of the injury consequences.
This thesis demonstrated that the three comorbidity adjustments methods to account for comorbidity effects (i.e. the maximum limit, additive and multiplicative approach) provided reasonable results. However, to strengthen the insight in current approaches to adjust for comorbidity, more research is needed on the effects of comorbidity on health-related quality of life measured with other, more sensitive MAUIs, such as the HUI.

3. Long-term sequelae

For burden of disease studies that aid priority-setting in health care and prevention, the burden of disease estimate should comprise a complete assessment of the health effects due to a certain cause. Nonetheless, a close investigation learned that marked sequelae are omitted whereas sequelae that are less likely to occur are included.
In this thesis two previously overlooked sequelae of infectious intestinal disease and injuries were investigated, i.e. post-infectious irritable bowel syndrome (PI-IBS) and posttraumatic stress disorder (PTSD), respectively. In case of PI-IBS, this revealed that it manifests itself in 9% of patients with bacterial infectious intestinal disease and that it accounts for half of the total disease burden for the selected pathogens. Regarding the

health domain injury, PTSD was shown to be relatively prevalent among injury patients. Shorter than 3 months after the injury a PTSD prevalence rate of 24% was found, tapering down to 4% at 24 months after the injury. Including PTSD increased the non-fatal burden of disease of unintentional injuries with 54%.
However, the examination of sequelae is complicated, because conspicuous difficulties may hamper the evidential value of the sequelae. Difficulties such as time delay between acute disorder and onset of the sequelae, and the fact that sequelae may be relatively common in the general population.
Nevertheless, for future burden of disease studies it is recommended to define all possible health outcomes of a certain injury or disease. These possible health outcomes should cover immediate as well as remote health outcomes and physical as well as psychological health outcomes. This implies that for all health domains, the definition of sequelae should be broadened to physical as well as psychological health outcomes.

4. Minimal disease

Related to the question regarding the question of in- and exclusion of sequelae is the question whether there should be a threshold, a cut-off point, reflecting a relevance criterion, which decides on the minimum of health impact accounted for in the calculation of DALYs. Below the threshold consequences are then assumed to be part of normal life, i.e. on average too trivial to account for. Such a cut-off point, or relevance criterion, may be based on health care use, anatomical features or health outcome.
This thesis showed that the preference-based relevance criterion, a criterion based on outcome, distinguishes experienced minimal disease from relevant disease unambiguously, yet it also allows relevant cases of minimal disease to be included in burden of disease estimates in order to avoid incomplete estimates. Application of this preference-based relevance criterion results in considerable changes in ranking of pathogens that may cause infectious intestinal disease.

Conclusions

This thesis demonstrated satisfactory alternative solutions to four controversial issues of the DALY concept. As a result, the DALY concept may be used for burden of disease studies of conditions in which these controversial issues emerge in particular. Application of alternative solutions proposed in this thesis may improve burden of disease estimates considerably.

Samenvatting

De volksgezondheid wordt bedreigd door een breed scala aan ziekten en letsels. Echter, voor bijvoorbeeld preventie en surveillance van deze ziekten en letsels zijn slechts beperkte middelen beschikbaar. Beleidsmakers zijn hierdoor genoodzaakt om zich te richten op de meest relevante ziekten en letsels. Om te bepalen welke ziekten en letsels het meest relevant zijn, ontwikkelden onderzoekers van de Wereldgezondheidsorganisatie (WHO) en de Wereldbank in 1993 een innovatief concept. Dit concept berekent hoeveel gezondheidsverlies in de bevolking door ziekten en letsels veroorzaakt wordt. De hoeveelheid gezondheidsverlies, oftewel ziektelast (´burden of disease`), wordt uitgedrukt in Disability Adjusted Life Years (DALYs). De DALY geeft in één getal weer hoeveel jaren verloren zijn door vroegtijdige sterfte en hoeveel jaren geleefd zijn met een ziekte of letsel. Door de ziektelast uit te drukken in DALYs kunnen onderzoekers en beleidsmakers de ziektelast van verschillende aandoeningen met elkaar vergelijken. Zo kunnen ze bijvoorbeeld de ziektelast van verkeerongevallen vergelijken met de ziektelast van maagdarminfecties. Op deze manier kan bepaald worden welke ziekten en letsels het meest relevant zijn in termen van ziektelast.

Het toepassen van het DALY-concept heeft dus grote voordelen. Het DALY-concept is echter ook bekritiseerd, omdat bepaalde methodologische aspecten van de DALY omstreden zijn. In dit proefschrift zijn vier van deze omstreden methodologische aspecten onderzocht met betrekking tot twee verschillende gezondheidsdomeinen, namelijk 1) ongevallen en letsels en 2) maagdarminfecties. Letsels en maagdarminfecties kenmerken zich door een grote variatie in gevolgen voor de gezondheid. Deze gevolgen kunnen licht en van slechts tijdelijke aard zijn, maar ook zeer ernstig en langdurig en deze variatie in gevolgen kan problemen opleveren bij het toepassen van het DALY-concept.

1. Wegingsfactoren en gezondheidsgerelateerde kwaliteit van leven

Het eerste omstreden methodologische aspect dat in dit proefschrift is onderzocht, betrof het proces om wegingsfactoren voor ziekten en letsels te bepalen. Deze wegingsfactoren zijn nodig om de tijd geleefd met ziekte of letsel te berekenen. De wegingsfactor drukt uit hoe ernstig de gevolgen van een aandoening zijn voor het functioneren van een patiënt en heeft een waarde tussen 0 (best voorstelbare gezondheidstoestand) en 1 (slechtst voorstelbare gezondheidstoestand).

Het oorspronkelijke DALY-concept vereist dat een gezondheidstoestand constant blijft voor een relatief lange periode. Dit is echter zeer problematisch voor aandoeningen die gekenmerkt worden door een complex verloop, zoals acute aandoeningen van korte duur. Daarnaast ontbraken wegingsfactoren voor veelvoorkomende lichamelijke letsels, waardoor de ziektelast van deze letsels niet berekend kon worden. In dit proefschrift zijn daarom 43 wegingsfactoren bepaald voor zowel langdurige als tijdelijke lichamelijke letsels. Om deze wegingsfactoren te bepalen werd de jaarprofielmethode gebruikt. Deze methode beschrijft de gezondheidstoestand van een patiënt voor een periode van één

jaar. Door een perspectief van één jaar te kiezen, is het mogelijk om ook wegingsfactoren te bepalen voor aandoeningen zich kenmerken door een complex verloop, zoals acute tijdelijke lichamelijke letsels. Een panel van 143 leken beoordeelde 43 letsels en op deze oordelen zijn vervolgens de wegingsfactoren gebaseerd. Door toepassing van deze 43 nieuwe wegingsfactoren nam ziektelast van letsels door ongevallen toe met 36%.

Wegingsfactoren kunnen op verschillende manieren worden bepaald. Eén manier is om een panel ziektespecifieke beschrijvingen van aandoeningen te laten beoordelen zoals hierboven is beschreven. De ziektespecifieke beschrijving geeft aan om welke aandoening het gaat en welke specifieke klachten de patiënt heeft. Een tweede manier om wegingsfactoren te bepalen, is door het gebruik van generieke (niet-ziektespecifieke) instrumenten, zoals het EQ-5D instrument. De EQ-5D meet de functionele beperkingen van een patiënt ten aanzien van mobiliteit, zelfverzorging, dagelijkse activiteiten, pijn of ongemak, en angst of somberheid. Hiermee kan de gezondheidsgerelateerde kwaliteit van leven van deze patiënt worden bepaald. De gegevens kunnen ook gebruikt worden om wegingsfactoren te bepalen.

In dit proefschrift zijn met de EQ-5D de functionele beperkingen gemeten van een populatie ongevalspatiënten die zijn behandeld op de spoedeisende hulpafdeling van een ziekenhuis. De patiënten vulden een vragenlijst in op vier verschillende tijdstippen, namelijk 2 ½, 5, 9 en 24 maanden na het ongeval. Deze vragenlijsten bevatten ook de EQ-5D. De EQ-5D gegevens van de patiënten werden vervolgens omgezet in wegingsfactoren.

Vergelijking van de generieke EQ-5D wegingsfactoren en de ziektespecifieke jaarprofielwegingsfactoren toonde aan dat de EQ-5D wegingsfactoren consequent hoger (´ernstiger´) waren vergeleken met de jaarprofielwegingsfactoren. Het verschil tussen de wegingsfactoren was vooral groot in geval van acute tijdelijke aandoeningen. Toepassing van de relatief hoge EQ-5D wegingsfactoren in ziektelastberekeningen van veelvoorkomende letsels kan daardoor resulteren in een overschatting van de ziektelast van deze letsels. EQ-5D wegingsfactoren lijken dan ook minder geschikt voor acute tijdelijke gezondheidstoestanden.

Vervolgens werd in dit proefschrift onderzocht of de beschrijving van de gezondheidstoestand invloed heeft op het oordeel van een panel leken. Om dit te onderzoeken, beoordeelden deelnemers van het lekenpanel een aantal gezondheidstoestanden tweemaal. Eenmaal werd de gezondheidstoestand beschreven met een ziektespecifieke beschrijving en eenmaal met een generiek beschrijving. Uit de oordelen bleek dat de wijze waarop de gezondheidstoestand wordt beschreven (ziektespecifiek versus generiek) grote invloed had op de beoordeling van de gezondheidstoestand. Dit verschil was vooral groot als het ging om lichte aandoeningen en aandoeningen waarbij het uiterlijk is aangetast, zoals huidaandoeningen. Als deze aandoeningen met een ziektespecifieke beschrijving werden gepresenteerd, beoordeelden de panelleden de aandoening als ernstiger. Dit resulteerde vervolgens in een hogere wegingsfactor. Hieruit bleek dat panelleden ziektespecifieke klachten belangrijk vinden bij het beoordelen van een gezondheidstoestand. Generieke

beschrijvingen bevatten geen informatie over ziektespecifieke klachten en dit lijdt vervolgens tot informatieverlies.
De hoeveelheid informatieverlies die optreedt bij het gebruik van een generiek instrument is sterk afhankelijk van het generieke instrument dat wordt gebruikt om de functionele beperkingen te meten. Bij patiënten met langdurige lichamelijke letsels werden de functionele beperkingen en de gezondheids-gerelateerde kwaliteit van leven gemeten met twee verschillende generieke instrumenten, de EQ-5D en de Health Utilities Index (HUI). De gegevens toonden aan dat de HUI een gevoeliger instrument is vergeleken met de EQ-5D. De EQ-5D meet bijvoorbeeld in sommige gevallen geen functionele beperkingen, terwijl de HUI wel beperkingen meet.

2. Comorbiditeit

Het tweede omstreden aspect dat in dit proefschrift werd onderzocht, is dat het DALY-concept geen rekening houdt met comorbiditeit. Er is sprake van comorbiditeit als een patiënt tegelijkertijd twee of meer stoornissen of aandoeningen heeft. Comorbiditeit is een veelvoorkomend probleem in de Nederlandse bevolking. Echter, tot op heden houden de meeste onderzoekers geen rekening met comorbiditeit bij het uitvoeren van ziektelastberekeningen. Hierdoor kunnen onderzoekers en beleidsmakers verkeerde conclusies trekken over de relevantie van bepaalde aandoeningen in termen van ziektelast ten aanzien van de volksgezondheid.
In dit proefschrift werden drie bestaande methoden onderzocht waarmee ziektelast-berekeningen kunnen worden bijgesteld voor comorbiditeit. Vanwege de tijdelijke aard van veel letsels zijn de EQ-5D gegevens van de ongevalspatiënten gebruikt om deze methoden te testen. De drie bestaande methoden bleken redelijk goede resultaten op te leveren, maar voordat deze methoden daadwerkelijk toegepast kunnen worden, is er meer onderzoek nodig.

3. Late gevolgen

Om de relevantie van ziekten en letsels te bepalen, is het belangrijk dat ziektelast-schattingen compleet zijn. Als bijvoorbeeld de ziektelastschatting van verkeersongevallen beperkt is tot acute lichamelijke letsels en langdurige letselgevolgen niet worden meegerekend, is er sprake van een incomplete schatting. Niettemin leert onderzoek dat relatief veelvoorkomende gevolgen en sequelae worden weggelaten, terwijl zeldzame gevolgen en sequelae wel worden meegenomen in de ziektelastberekeningen. Een sequela is een aandoening die het gevolg is van een ziekte of letsel.
In dit proefschrift werden twee sequelae bestudeerd die eerder niet werden opgenomen in ziektelastberekeningen, namelijk posttraumatische stress-stoornis (sequela van een ongeval) en postinfectieuze prikkelbare darmsyndroom (sequela van een maagdarm-infectie). Het onderzoek naar het post-infectieuze prikkelbare darmsyndroom toonde aan dat bijna één op de tien patiënten met een maagdarminfectie daarna aan het postinfectieuze prikkelbare darmsyndroom lijdt. Als het postinfectieuze prikkelbare

darmsyndroom wordt opgenomen in de ziektelastberekeningen van maagdarminfecties, dan neemt de ziektelast van maagdarminfecties met meer dan 50% toe. Het onderzoek naar de sequela posttraumatische stress-stoornis liet zien dat deze stoornis veelvuldig voorkomt bij ongevalspatiënten behandeld op de spoedeisende hulpafdeling of opgenomen in het ziekenhuis. In de eerste drie maanden na het ongeval bleek dat één op de vier patiënten een posttraumatische stress-stoornis had. Vierentwintig maanden na het ongeval was dit aantal afgenomen tot minder dan één op de twintig. Daarnaast bleek dat als posttraumatische stress-stoornis wordt meegerekend, de ziektelast van ongevallen toeneemt met 54%.
Het onderzoek naar sequelae wordt bemoeilijkt doordat er veel tijd kan zitten tussen de acute aandoening en het ontwikkelen van de sequela. Bovendien kunnen bepaalde sequelae ook andere oorzaken hebben en/of veel voorkomen in de bevolking. Dit maakt het lastig om te bepalen of er sprake is van een sequela, of dat de aandoening een andere oorzaak had. Desalniettemin is het bij het berekenen van de ziektelast belangrijk om alle mogelijke directe en later optredende gevolgen van een aandoening in kaart te brengen en op te nemen in de ziektelastberekeningen.

4. Minimale ziekte

Het vierde en laatste controversiële aspect van het DALY-concept dat in dit proefschrift is onderzocht, is het relevantie criterium. Met het relevantie criterium kan worden bepaald of een aandoening te triviaal is om mee te nemen in de ziektelastberekeningen. Als een aandoening niet voldoet aan het relevantie criterium, dan wordt de aandoening uitgesloten. In het oorspronkelijke DALY-concept wordt er geen drempel voor triviale gezondheidstoestanden toegepast. Dit beperkt de toepassing van het DALY-concept bij ziektelast-berekeningen van aandoeningen met lichte en/of tijdelijke gevolgen.
Een relevantie criterium kan gebaseerd zijn op gebruik van de gezondheidszorg, zoals bijvoorbeeld huisartsbezoek, anatomische kenmerken of het verwachte effect op de gezondheidstoestand. Dit proefschrift toonde aan dat het relevantie criterium dat gebaseerd is op het verwachte effect op de gezondheidstoestand onderscheid maakt tussen triviale en relevante aandoeningen, maar aandoeningen die minimale functionele beperkingen veroorzaken en weinig effect op de gezondheidstoestand hebben, worden niet uitgesloten. Dit voorkomt een incomplete ziektelastschatting. Daarnaast liet dit proefschrift zien dat de toepassing van het relevantie criterium bij de ziektelast-berekeningen van maagdarminfecties de rangorde, en daarmee de relevantie, van de verschillende ziekteverwekkers die maagdarminfecties veroorzaken aanzienlijk verandert.

Conclusies

Met het DALY-concept, ontwikkeld door de WHO en de Wereldbank, kan bepaald worden welke ziekten en letsels de meest relevante bedreigingen vormen voor de volksgezondheid. Verschillende aspecten van het DALY-concept zijn echter bekritiseerd. Bovendien veroorzaakten deze omstreden aspecten problemen bij het berekenen van de

ziektelast van bepaalde aandoeningen, zoals letsels en maagdarminfecties. In dit proefschrift zijn vier controversiële methodologische aspecten van het DALY-concept en alternatieve oplossingen onderzocht. Hieruit bleek dat toepassing van de voorgestelde alternatieve oplossingen de ziektelastberekeningen van deze aandoeningen aanzienlijk kunnen verbeteren.

Dankwoord

Tot slot wil ik graag iedereen bedanken met wiens hulp dit proefschrift tot stand is gekomen.

In de eerste plaats bedank ik mijn promotor Prof.dr. G.J. Bonsel voor zijn vertrouwen in mij en de ruimte die hij mij gegeven heeft. Beste Gouke, jouw scherpe blik en vermogen om onderzoek naar een hoger plan te tillen zijn bewonderenswaardig. Veel heb ik van jou geleerd, zoals het feit dat overleggen met stroopwafels zoveel smakelijker is.
Daarnaast wil ik mijn co-promotor Dr. E.F. van Beeck bedanken voor zijn steun en begeleiding. Beste Ed, ik waardeer het enorm dat jouw deur altijd voor me open staat. Mijn ideeën en artikelen werden door jou met onverminderd groot enthousiasme begroet. Graag wil ik je hiervoor bedanken, alsmede voor de wijze raad die je me in de afgelopen jaren gegeven hebt.

Prof.dr. J.J. van Busschbach, Prof.dr. P. Patka en Prof.dr. A.H. Havelaar wil ik bedanken voor de tijd en moeite die ze hebben genomen om mijn proefschrift te beoordelen.
Tevens wil ik Prof.dr. A.H. Havelaar bedanken voor de prettige samenwerking. Beste Arie, met veel plezier heb ik met jou samengewerkt bij het RIVM en geschreven aan de artikelen die uit deze samenwerking voortgekomen zijn, mede dankzij jouw efficiëntie en vermogen om ingewikkelde dingen begrijpelijk uit te leggen. Bedankt hiervoor.

Beste Suzanne, jou wil ik ook hartelijk bedanken voor je hulp en steun bij het uitvoeren van mijn promotieonderzoek. Ik heb bewondering voor je tomeloze inzet en de manier waarop jij gezelligheid en hard werken combineert.
Veel andere collega's hebben bijgedragen aan het uitvoeren van mijn onderzoek en het schrijven van mijn proefschrift, waarvoor ik iedereen heel hartelijk dank.
Daarnaast wil ik mijn collega´s van de afdeling Sociale Geneeskunde van het AMC, de afdeling LZO van het RIVM en de afdeling Maatschappelijke Gezondheidszorg van het EMC en in het bijzonder mijn kamergenoten Tanja, Bas, Denise, Jack, Katsuhisha, Martijn (nu met discobal) en Willemijn, bedanken voor de gezellige sfeer waarin ik aan mijn promotieonderzoek heb gewerkt.
Mijn nieuwe collega´s bij Acute Zorgregio Oost wil ik bedanken voor alle interesse in de vorderingen omtrent mijn proefschrift.

Zonder panels en letselpatiënten zouden er geen gegevens zijn geweest om te onderzoeken. Daarom wil ik de deelnemers aan het IBIS en PsyBIS onderzoek en de patiënten die meededen aan de patiëntenquêtes bedanken voor hun deelname. De onderzoekers van Stichting Consument en Veiligheid, die de gegevens van de letsel-patiënten verzameld hebben, bedank ik voor hun samenwerking

Ook buiten de onderzoekswereld hebben een heel aantal mensen bijgedragen aan het prettige verloop van mijn promotieonderzoek.
Beste vrienden, familie en Nesbewoners, bedankt voor jullie belangstelling voor mijn onderzoek.
Lieve Dorine, weer een belangrijke dag waarop jij aan mijn zijde staat. Bedankt voor alle keren dat jij het leven weer in perspectief plaatst.
Lieve Gerda en Olle, Iwan en Mariska, bedankt voor de gezelligheid die met jullie bezoekjes gepaard gaan en de veelal broodnodige afleiding.
Leave heit en mem, tige tankbar bin ik foar jim nea ôflittende steun en de wittenskip dat ik altiid bij jimme terjucht kin, foar praktiske hulp of in lústerjend ear.
Lieve Sjors, in de afgelopen jaren heb jij meerdere malen delen van mijn proefschrift teruggehaald uit de digitale Bermudadriehoek en mijn computer gereanimeerd. Maar meer nog ben ik dankbaar en blij dat jij vandaag naast me staat.

Curriculum Vitae

Juanita Haagsma werd op 3 september 1976 geboren te Heerenveen. Na de MAVO en de HAVO, behaalde ze in 1996 haar VWO eindexamen. Na een jaar bewegingswetenschappen, besloot ze psychologie te gaan studeren aan de Vrije Universiteit en in 2001 studeerde ze af in de richting biologische psychologie.
Van begin 2002 tot september 2003 werkte ze als data-analist bij Stichting Consument en Veiligheid en werd haar belangstelling gewekt voor het meten van de ziektelast van letsels. Vanaf 2003 werkte ze onder begeleiding van haar promotor Prof.dr. G.J. Bonsel en haar co-promotor Dr. E.F. van Beeck aan haar promotieonderzoek naar de methodologie van samengestelde volksgezondheidsmaten. Dit onderzoek werd uitgevoerd op de afdeling Sociale Geneeskunde van het AMC te Amsterdam. In 2007 vervolgde ze het promotieonderzoek in deeltijd aan de Erasmus Universiteit en daarnaast werkte ze bij het Laboratorium voor Zoönosen en Omgevingsmicrobiologie (LZO) van het Centrum voor Infectieziekten Bestrijding van het RIVM, alwaar zij onder begeleiding van Prof.dr. A.H. Havelaar onderzoek uitvoerde naar de ziektelast van voedseloverdraagbare aandoeningen, waaronder maagdarminfecties.
Momenteel werkt Juanita bij de afdeling Maatschappelijke Gezondheidszorg van het ErasmusMC en voert ze verder onderzoek uit naar de ziektelast van letsels en de methodologie van samengestelde volksgezondheidsmaten. Daarnaast werkt ze als onderzoeker bij Acute Zorgregio Oost van het UMC St Radboud.
Juanita is getrouwd met Sjors Rozendaal.

List of publications

Publications and manuscripts printed in this thesis

Haagsma JA, van Beeck EF, Polinder S, Hoeymans N, Mulder S, Bonsel GJ. Novel empirical disability weights to assess the burden of non-fatal injury. Injury Prevention 2008; 14(1): p. 5-10.

Haagsma JA, Polinder S, van Beeck EF, Mulder S, Bonsel GJ. Alternative approaches to derive disability weights in injuries: do they make a difference? Quality of Life Research 2009; 18(5): p. 657-65.

Haagsma JA, Janssen MF, Birnie E, Bonsel GJ. Comparison of generic and disease specific health state valuations by a laymen panel. Conference proceedings of the EuroQol Plenary Meeting 2005.

Polinder S, Haagsma JA, Bonsel GJ, Essink-Bot ML, Toet H, van Beeck EF. The measurement of long-term health-related quality of life after injury: comparison of the EQ-5D and the Health Utilities Index. Injury Prevention. 2010; 16(3): p. 147-53.

Haagsma JA, van Beeck EF, Polinder S, Toet H, Bonsel GJ. Independent comorbidity and health-related quality of life of injury patients in the first year following injury: comparison of three adjustment approaches. Submitted.

Haagsma JA, Siersema PD, de Wit NJ, Havelaar AH. Disease burden of post-infectious irritable bowel syndrome. Epidemiology and Infection 2010; 138(11): p.1650-6.

Haagsma JA, van Beeck EF, Polinder S, Olff M, Mulder S, Bonsel GJ. Posttraumatic stress symptoms and health-related quality of life two years after injury treated at the Emergency Department. Submitted.

Haagsma JA, Polinder S, Toet H, Panneman M, Havelaar AH, Bonsel GJ, van Beeck EF. Beyond the neglect of psychological consequences: posttraumatic stress disorder increases the non-fatal burden of injury by more than 50%. Accepted by Injury Prevention.

Polinder S, Meerding WJ, Lyons RA, Haagsma JA, Toet H, Petridou ET, Mulder S, van Beeck EF. International variation in clinical injury incidence: exploring the performance of indicators based on health care, anatomical and outcome criteria.
Accident, Analysis and Prevention 2008; 40(1): 182-91.

Haagsma JA, Havelaar AH, Janssen MF, GJ Bonsel. Disability Adjusted Life Years and minimal disease: application of a preference-based relevance criterion to rank enteric pathogens. Population Health Metrics 2008; 6(1): 7.

Other publications and manuscripts

Polinder S, Haagsma JA, Toet H, Brugmans MJP, van Beeck EF and the EUROCOST and APOLLO reference groups. Burden of injury in childhood and adolescence in Europe: time for action now. BMC Public Health 2010; 10: p. 45.

Haagsma JA, van der Zanden BP, Tariq L, van Pelt W, van Duynhoven YTPH, Havelaar AH. Disease burden and costs of selected foodborne pathogens in the Netherlands, 2006. Bilthoven, National Institute for Public Health and the Environment, 2009.

Haagsma JA, Belt E, Polinder S, Lund J, Atkinson M, Macey S, Lyons, RA, van Beeck EF. INTEGRIS WP5 report: injury disability indicators. Department of Public Health, ErasmusMC, Rotterdam, 2009.

Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the standard EQ-5D three-level system with a five-level version. Value in Health 2008; 11(2): p. 275-84.

Haagsma JA, Bonsel GJ. [Disability weights for foodborne disease.] Department of Social Medicine, AMC, Amsterdam, 2007.

Haagsma JA, Gommer AM, Bonsel GJ. [Disability weights for occupational-related disorders.] Department of Social Medicine, AMC, Amsterdam, 2007.

Haagsma JA, Bonsel GJ. [Disability weights for posttraumatic stress disorder.] Department of Social Medicine, AMC, Amsterdam, 2007.

Haagsma JA, Janssen MF, Beeck E van, Hoeymans N, Mulder S & Bonsel GJ. The Integrated Burden of Injury Study. Department of Social Medicine, AMC, Amsterdam, 2007.

Haagsma JA, Janssen MF, Birnie E, Bonsel GJ. Health state valuation of Attention Deficit Hyperactivity Disorder. Amsterdam: Department of Social Medicine, Academic Medical Center/University of Amsterdam, 2005.

Haagsma JA, Janssen MF, Bonsel GJ. [Disability weights for gastroenteritis, hepatitis A and ocular toxoplasmosis. A contribution to the national foodborne disease estimate.] Department of Social Medicine, AMC, Amsterdam, 2007.