UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Disability weights for diseases in the Netherlands

Stouthard, M.E.A.; Essink-Bot, M.L.; Bonsel, G.J.; Barendregt, J.J.M.; Kramers, P.G.N.; van de Water, H.P.A.; Gunning-Schepers, L.J.; van der Maas, P.J.

Publication date
1997

Document Version
Final published version

Link to publication

Citation for published version (APA):
Stouthard, M. E. A., Essink-Bot, M. L., Bonsel, G. J., Barendregt, J. J. M., Kramers, P. G. N., van de Water, H. P. A., Gunning-Schepers, L. J., & van der Maas, P. J. (1997). *Disability weights for diseases in the Netherlands*. Inst. Sociale Geneeskunde.

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:19 Apr 2021

DISABILITY WEIGHTS FOR DISEASES IN THE NETHERLANDS

Cover design: Paul F.M. Krabbe, Theta Research, Zeist
Layout: Bon Mot, Rotterdam

ISBN: 90-72245-84-9

Subject headings: disability, disability weights, composite health outcome measures, DALY, QALY, burden of disease, valuation methods

© 1997, Department of Public Health, Erasmus University Rotterdam, the Netherlands
All rights reserved.
No part of this publication may be reproduced in any form by print, photoprint, microfilm or any other means without written permission of the rightful claimant(s). This restriction concerns the entire publication or any part of it.

Correspondence for reproduction and ordering:
Department of Public Health, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, the Netherlands, fax +31 10 436 6831, e-mail MGZ@MGZ.FGG.EUR.NL

The project 'Disability weights for diseases in the Netherlands' was funded by the Dutch Ministry of Health, Welfare and Sports.

DISABILITY WEIGHTS FOR DISEASES IN THE NETHERLANDS

Marlies E.A. Stouthard
Marie-Louise Essink-Bot
Gouke J. Bonsel
Jan J. Barendregt
Pieter G.N. Kramers
Harry PA van de Water
Louise J. Gunning-Schepers
Paul J. van der Maas

Institute of Social Medicine, Academic Medical Centre / University of Amsterdam
Department of Public Health, Erasmus University Rotterdam
Institute of Clinical Epidemiology and Biostatistics, Academic Medical Centre / University of Amsterdam
Public Health Status and Forecast Centre, National Institute of Public Health and the Environment, Bilthoven
TNO Prevention & Health, Leiden

The Netherlands

PREFACE

ANY COMPREHENSIVE measure for the burden of disease in a population has to contain information about morbidity and mortality. Depending on a number of choices, such as whether prevalences are used or incidence-duration-mortality estimates, whether life table populations are used or dynamic demographic models, a whole 'family of measures' can be derived. Examples are the QALY, the Healthy Life Expectancy, and the DALY. The latter stems from the Global Burden of Disease project, which was a major breakthrough in its rigorous application of uniform methodology.

The next steps to be taken in this area are the further refining of the methodology and finding a balance between making the information sufficiently country- or region-specific, while at the same time maintaining comparability between countries. Such information should contribute to a more rational health policy making, also in the sense that it generates reference data that are indispensable for the economic evaluation of any health intervention.

The present report is one step towards such a burden-of-disease framework for the Netherlands. It describes disability weights for the Netherlands and how they were derived, building further on the work and with the support of Christopher Murray. The study was funded by the Dutch Ministry of Health, Welfare and Sports. Further improvement and validation of the methodology to derive the weights, their international transferability and their applicability will be investigated in an international context. A European research network consisting of groups in Great Britain, France, Spain, Sweden, Norway, Denmark and the Netherlands is being funded by the EU within the scope of the BIOMED-II program. It will launch an international study (starting 1998) into the similarities and differences between these countries with respect to disability weights for diseases and the possibilities for comparable country-specific burden-of-disease estimates.

Prof.dr Paul J. van der Maas
Rotterdam, December 1997

SUMMARY

IN THE project on 'Disability Weights for Diseases', a coherent set of disability weights was derived for a sizable number of diseases. In principle, it consequently became possible to combine, in a comparable manner, data on mortality and the functional sequelae for all these diseases into a single measure. Public health research and the economic evaluation of health care interventions offer important application possibilities for the disability weights derived. The 'Disability weights for diseases' project was conducted further to the Public Health Status and Forecast [report for 1997] (*VTV-97)*, as a collaborative project between the Institute of Social Medicine, Academic Medical Centre / University of Amsterdam; the Department of Public Health, Erasmus University Rotterdam; TNO-Prevention & Health, Leiden; and the National Institute of Public Health and the Environment, Bilthoven, all in the Netherlands.

Chapter 1 explains why disability weights offer a valuable addition to the available arsenal of measures for the health of populations. The Global Burden of Disease study (GBD) performed by Murray and Lopez at the request of the Worldbank and WHO is described as an illustration. Extensive attention is devoted to the methodology used in the GBD study for deriving the disability weights. Finally, the objectives of the Dutch project are described.

In chapter 2, the design of the Dutch disability weights study is described. The set-up corresponds to that of the GBD study, with some amendments. The adapted protocol was tested in a pilot study. The list of diseases for which disability weights were derived was taken from the Public Health Status and Forecast 1997 study (*VTV-97)* and comprised for the present study a total of 53 diagnostic groups. Each diagnostic group was broken down into one or more homogenous disease stages according to health status, treatment and prognosis. All disease stages were provided with a representative description of the functional health state in terms of an extended version of the EuroQol 5D classification. The 175 disease stages on the final list were submitted, in accordance with the protocol, to a number of expert panels in order to enable disability weights to be derived. This occurred in two steps. First, weights were derived as meticulously as possible by the panels in an interactive procedure using the person trade-off (PTO) method and a visual analogue scale (VAS) for 16 indicator conditions. The disability scale was calibrated by the positioning of these 16 weights. The weights for the remaining disease stages were subsequently elicited with the help of a written interpolation procedure on the basis of the calibrated disability scale. At the end of chapter 2, the method followed for investigating the reliability and validity of the weights is discussed.

In chapter 3, the results of the study are presented. The disability weights protocol for the panel sessions proved to be relatively easy to implement. The 16 indicator conditions were spread across the entire disability scale. The interpanel reliability of the values on the scale proved to be satisfactory. The reliability of the interpolations was calculated for six 'common core' disease stages, which were interpolated by all of the experts. The agreement between the panel members and the test-retest reliability at group level was good. The test-retest reliability at the level of the individual was moderate. The validity of the disability weights was bolstered by comparison with the GBD disability weights, the comparison in and between diagnostic groups and by comparison with the weights estimated with the help of a theoretical model based on the extended EuroQol description.

Chapter 4 contains the conclusions and recommendations of the study. It may be concluded that a coherent set of disability weights was elicited in a reliable way for a large number of diseases on the basis of the Dutch protocol. Application of this set of weights in economic evaluations will promote the mutual commensurability of these studies. Simultaneous application of these weights in economic evaluation studies and public health research will foster the integration of information from both these areas. Further research into the reliability and validity of the disability weights derived is recommended. The usability of the weights in public health research will primarily depend on the availability of consistent and comprehensive epidemiological data on the relevant conditions and disease stages. Further research into the representativeness and the accuracy of the standardized description of the functional health state for each disease stage is also demanded. This may require refinement of the classification used to describe the functional health states. The present disability weights may possibly be too crude for application in evaluation studies of specific health interventions. Disaggregation of the weights for specific diseases and disease stages using the same methods of determination can yield the desired refinement while retaining commensurability with the existing scale. Finally, further research is recommended on trends in disability weights seen over the course of time and resulting from e.g. new treatment methods, on weights for combinations of diseases (co-morbidity) and on international comparisons and application of the disability weights.

CONTENTS

1	**INTRODUCTION**	1
1.1	**Composite health outcome measures: more than mortality alone**	1
1.1.1	Combining mortality and morbidity	2
1.1.2	The common denominator: time	3
1.1.3	Weights in medical evaluation research: QALYs	3
1.1.4	The Global Burden of Disease Study	4
1.2	**Determining the weights in the Global Burden of Disease Study**	5
1.2.1	Valuation method: Person Trade-Off	6
1.2.2	Respondents in the GBD study: medical experts	7
1.2.3	The valuation procedure in the GBD study: choice of indicator conditions	8
1.3	**Objectives of the Dutch disability weights project**	9
2	**DESIGN OF THE STUDY ON DISABILITY WEIGHTS FOR DISEASES IN THE NETHERLANDS**	11
2.1	**Pilot study**	12
2.2	**The list of diseases and disease stages**	13
2.2.1	The diseases	13
2.2.2	Disease stages	13
2.2.3	Addition of standardized health state description to disease stage	14
2.2.4	Duration	15
2.2.5	The final lists of disease stages	16
2.3	**Selection of the panel members**	17
2.4	**Valuation methods: PTO, VAS**	22
2.5	**Selection of the indicator conditions**	22
2.6	**Description panel session**	27
2.7	**Interpolation**	28
2.8	**Test-retest**	30
2.9	**Analysis of the data**	30
2.9.1	Calculating the disability weights	30
2.9.2	Reliability and validity of the weights	31

3	**RESULTS: DISABILITY WEIGHTS FOR DISEASES**	33
3.1	**Description of the panels**	33
3.2	**Results of the panel sessions**	35
3.2.1	Disability weights for the indicator conditions	36
3.2.2	Reliability of the values found for the indicator conditions	37
3.2.3	Disability scale	37
3.3	**Results of the interpolation session**	40
3.3.1	Interpolation	40
3.3.2	Realiability of the interpolations	40
3.4	**Validity of the weights**	47
3.4.1	Comparison with the weights from the GBD study	49
3.4.2	Comparison of disability weights per disease	49
3.4.3	Comparisons of disability weights between diseases	50
3.4.4.	Assessing validity of the weights using EuroQol 5D+ classifications	52
3.5	**Results lay panel**	54
4	**CONCLUSIONS AND RECOMMENDATIONS**	59
4.1	**Possible uses**	59
4.2	**Research recommendations**	61
4.2.1	The need for corresponding epidemiological data	61
4.2.2	Standardized description of each disease stage	61
4.2.3	Reliability and validity of the current disability weights	62
4.2.4	Trends in disability weights	62
4.2.5	Disaggregation of disability weights for application in assessment studies of spcific interventions	63
4.2.6	Co-morbidity	63
4.2.7	International comparisons and application	63
	REFERENCES	65
	APPENDIX	
	Table A1 – Diagnostic groups, disease stages and EuroQol 5D+ classifications; incl. disability weights with 95% confidence intervals	67

1

INTRODUCTION

A COHERENT SET of disability weights was elicited for a large number of diseases in the project on 'Disability Weights for Diseases in the Netherlands'. This consequently enables the data on mortality and functional sequelae for all these diseases in principle to be combined, in a comparable manner, into a single measure. This report offers a detailed look at the design, implementation and outcome of this project.

This introduction outlines the background of the project. First, it examines how disability weights can constitute a meaningful addition to the available arsenal of methods for measuring the state of health of entire populations. Next, the method used by Murray and Lopez in the WHO/Worldbank 'Global Burden of Disease' study for determining disability weights for diseases is discussed. Finally, the goals of the Dutch disability weights project are specified.

1.1 Composite health outcome measures: more than mortality alone

A plethora of indicators are available by which to represent the state of a population's health. An important and fundamental distinction should be recognized between indicators based on mortality and indicators based on morbidity. Mortality-based indicators, for example, are the (whether or not disease-specific) mortality figures and a set of derived measures such as various standardized measures of mortality, premature mortality (in a number of variants) and conversely, the potential life years able to be gained, and life expectancy at birth. Other indicators are based on morbidity, such as, for example, the figures for new and existing cases of specific diseases and their sequelae in terms of disabilities and handicaps.

1.1.1 Combining mortality and morbidity

Morbidity and mortality are complementary aspects of the population's health. Obviously, as the ideal is a long life in good health, a good measure of the population's health should comprise both aspects. The advantage to a

mortality-based method of measuring the population's health is the fact that mortality is more easily measured than morbidity and that a (virtually) full-fledged registration system is in place. This indicator of measuring the population's health is therefore readily available and reliable. By contrast, morbidity is far more difficult to classify and to measure, while registration tends to be incomplete, unreliable or simply absent altogether.

Simplicity and reliability, therefore, argue in favour of a mortality-based measurement of the population's health. And if the two aspects of the population's health, morbidity and mortality could be summarized into a single mortality-based measure, the choice would be a straightforward one. However, only diseases running an acute course tend to be able to fulfil this condition: the patient falls ill and is either dead or has recovered without any notable consequences within a relatively brief period of time. Examples of such diseases are infectious diseases such as cholera and pneumonia. Other disorders are more chronic in nature (or leave lasting effects, such as poliomyelitis), and are therefore poorly captured solely by measuring mortality. Examples include conditions affecting the musculoskeletal system, various types of cancer, non acutely fatal cardiovascular disease, psychiatric disorders and dementia. Generally speaking, conditions leading to speedy fatality cause little morbidity (but possibly a high rate of mortality), while chronic diseases cause a high rate of morbidity (and possibly also, but not necessarily, high mortality).

Hence some sort of rift may be discerned between morbidity and mortality, between conditions which primarily cause morbidity and those primarily resulting in death. Obviously, therefore, an indicator based on morbidity would yield an entirely different view of the health problems within a population than would a mortality-based indicator. Yet this difference is also demonstrated by the developments within a single disease. Since the eighties, for example, the number of deaths due to heart disease in the Netherlands has been falling (improvement) while the number of existing cases of ischemic heart disease (deterioration) has concomitantly increased, and in particular, a more severe form has been on the rise (deterioration). This can be explained as follows: improving the survival rate for acute myocardial infarctions has yielded more heart patients, who subsequently run a considerable risk of developing heart failure. A public health measure based on mortality figures alone will mark the improvement (declining mortality) but not the change for the worse (more and more severe illness) following from this development, and hence presents a distorted view of reality.

The solution is, in principle, straightforward: we create a public health indicator encompassing both the aspects of morbidity and mortality. Such an indicator, combining morbidity and mortality is called a Composite Health Outcome Measure, or CHOM. Examples of CHOMs are:

- 'Healthy Life Expectancy' (in different variants, such as Life Expectancy in Perceived Good health and Life Expectancy without Disability)
- 'Disability Adjusted Life Expectancy' (DALE)
- 'Disability Adjusted Life years' (DALYs).

- ‘Quality Adjusted Life years’ (QALYs)

1.1.2 The common denominator: time

In a CHOM, morbidity and mortality must somehow be combined under a common denominator. Successful indicators based on mortality, such as life expectancy and lost life-years, use ‘time’ as the unit of measurement for the population’s health indicator, and more specifically, the number of years lived (life expectancy) or by contrast, not lived (the lost life-years). In combining morbidity and mortality into a single composite measure, the obvious choice is to express morbidity in terms of time as well: the years lived with disease. By rendering the time lost due to disease equivalent to the time lost through death it is possible to construct a CHOM. This is done by partly equating the time lived with disease with the time lost due to death. With the help of a disability weight, which reflects the relative severity of the disease, part of the time lived with the disease is regarded as not lived, and the remainder is regarded as time lived in good health.

In determining the disability weights, the severity of the condition is assessed at the level of the physical, mental and social functioning of the patient. On a scale of 1.00 (‘extreme functional consequences’) to 0.00 (‘no adverse functional effects at all’), a cold, for example, could be assigned a disability weight of 0.01 and multiple sclerosis of 0.67. Not lived years are assigned the weight 1.00 (‘dead’). A year lived with a disease assigned a weight of 0.40 will yield 0.60 ‘healthy life-year equivalents’ compared to 1.00 ‘healthy life-year equivalent’ for a full year of good health and 0.00 ‘healthy life-year equivalents’ for a year not lived due to premature death. By multiplying the time lived (in years) by the weight assigned to the state in which the years were spent, it becomes possible to compare both the functional effects of various diseases. Moreover, it becomes possible to compare the consequences of morbidity and mortality .

1.1.3 Weights in medical evaluation research: QALYs

Weighting the unhealthy years lived according to the degree of dysfunction in which they are spent has been customary in the economic evaluation (cost effectiveness or better cost utility analysis) of medical interventions. The consequences of morbidity and the changes therein are combined with data on survival in Quality Adjusted Life years (QALYs). An important problem in using the results of cost effectiveness analyses for policy objectives is the lack of uniformity in the methods used for deriving QALYs. Uniformity should be sought on at least four key points:

- the classification in describing the conditions to be evaluated
- the valuation method applied
- the choice of participants to perform the assessment
- the manner in which weights and years lived are combined into QALYs.

The value of QALY calculation for policy decisions on a public health level has therefore hitherto remained modest. The Council for Health and Social

Service (*Raad voor Volksgezondheid en Zorg*) in the Netherlands concluded in a recent report that a cost effectiveness methodology constitutes in principle a usable and available instrument for drug assessment in deciding whether or not a particular drug should be included in the health insurance package insured. The council, however, deems it vital that guidelines for conducting cost effectiveness analyses be drafted. One of the bottlenecks is how to quantify and value the results, or in other words, the way in which to generate QALYs. (Raad voor de Volksgezondheid en Zorg, 1997) The report in question can offer a contribution to the discussion ultimately aimed at compiling such guidelines.

1.1.4 The Global Burden of Disease Study

Recently an important study, the Global Burden of Disease (GBD) study was carried out at the request of the Worldbank and the WHO, by Murray (Harvard University) and Lopez (WHO). The first version of this study appeared as a Worldbank report in 1993, while a series of books on this project started appearing mid 1996. (Worldbank, 1993; Murray, 1996a) The GBD study is exceptional because of its worldwide scope. Data on mortality and the incidence/prevalence of diseases were collected for eight regions. In addition, a set of coherent weights was derived for the functional sequelae of a large number of diseases and injuries due to accidents. By combining both of these into a composite public health measure, it became possible to estimate the total burden of disease at the global level, and to break this down according to the share accounted for by specific diseases. Unique in the derivation of the weights in the GBD study is that all the steps and choices made throughout the process are transparent and substantiated by argumentation. The weights can be, and are, used for various applications, such as for describing regional patterns of ‘disability adjusted life expectancy’ (DALE) and for ascribing the burden of disease [expressed in disability adjusted life-years (DALYs)] to different causes.(Murray, 1997a; 1997b)

The GBD study has demonstrated the potential value of combining data about length of life and severity of disease in a single comprehensive measure. Descriptions of the population’s health with the help of such a measure may serve as a source of information for public health policy and for prioritizing and planning health care and health services research. The simultaneous application of such weights in health services research, including economic evaluation in health care can contribute importantly to the integration of information in both these areas. In the 1993 Worldbank report this is evidenced, for example, by the selection of essential packages of clinical facilities based on the scope and distribution of the health problems (expressed in terms of lost DALYs). The results of economic evaluation studies on drugs and other health care facilities in which the disability weights were applied may be used to establish priorities within and across categories of health care facilities, as recommended inter alia by the

Scientific Council for Government Policy (*Wetenschappelijke Raad voor het Regeringsbeleid).*(WRR, 1997)

The following section will explore extensively the methods used in the GBD study to elicit the weights for disease sequelae.

1.2 Determining the weights in the Global Burden of Disease study

The four key points recur in the method applied in the GBD study to elicit disease-specific disability weights (*see section 1.1.3*).

First, the classification of diseases occurred in a 'naturalistic' manner, i.e. the diseases to be valued were described on the basis of diagnostic labels and not as generic health states.

A second key point to the Murray method was that the valuation was set up as a two-step procedure. The first step involved the calibration of a disability scale. The weights were determined by positioning some 22 so-called indicator conditions on the scale. On the basis of this calibration, the scale was divided into seven more or less homogenous classes. During the second step, a huge number of other conditions were assigned to these classes. The relatively unknown valuation method of 'person trade-off' was applied to estimate the disability weights for the indicator conditions.

The third key point was that the weights were assigned by panels of medical experts.

A fourth point was the specific data processing to arrive at Disability Adjusted Life years (DALYs). The 'years lived with disability' were calculated by weighting one year periods for the health state in which these periods were spent through multiplication with disability weights. These were added to the 'years of life lost due to premature death'. The final step in the procedure to arrive at DALYs in the GBD study included additional adjustment for age (age weighting) together with a time preference (discounting). Age weighting and discounting are by no means essential to the DALY concept.

Various aspects of the procedure are examined in closer detail in the following.

1.2.1 Valuation method: Person Trade-Off

The person trade-off (PTO) method is a means of evaluating health states according to relative severity. An individual is asked to trade off healthy person-years and person-years lived with a disability. This method was first described by Patrick, Bush and Chen. (Patrick, 1973) Very few applications of this valuation method are found in the literature. In 1995, the method was retrieved from obscurity by Nord, in a publication in Medical Decision Making.(Nord, 1995)

Broadly speaking, there are three types of methods to value health states:

- methods in which health states are rated according to a given scale, known as direct rating. An example is the 'visual analogue scale' (VAS).
- 'trade-off' methods, in which individuals making the assessment are asked to surrender something in exchange for an improvement in health state. In the 'standard gamble' (SG), the person making the assessment must imaginarily trade off the certainty of surviving in a sub-optimum condition to gain an improvement in health state. The more (hypothetical) improvement in health the individual can achieve, the more (hypothetical) risk of death he will be willing to take. The time trade-off (TTO) allows an individual to trade-off an (hypothetical) improvement in health state for a (hypothetical) contraction of the quantity of life years.
- equivalence methods, such as for example 'equivalence of numbers', in which it is asked what number of persons in the health state to be assessed is equivalent to a number *x* in good health; and, for example, Olsen's method of operationalizing person trade-off, in which an equivalence is sought between an improvement in health state distributed on the one hand among more people for a shorter period of time and on the other hand among fewer people, but for longer. (Olsen, 1994)

Person trade-off is derived from the 'equivalence of numbers' method. Person trade-off differs from other trade-off methods in that individuals are asked to state preferences *between* people instead of *within a single* person. The use of PTO to assess states of health purports, more than other methods, to correspond with the perspective of a policymaker. This method is therefore possibly more suitable than, for example, VAS, SG and TTO, for estimating health state values intended for use at macro level (public health policy).(Pinto Prades, 1997) The equivalence method in principle requires no trade-off. However, since the very first publication on the equivalence of numbers technique, this has been operationalized as a trade-off. The study reported by Ubel is another example of an application of PTO. (Ubel, 1996)

In the GBD study, two variants of PTO were applied. In the first, PTO1, a respondent is asked to decide how for many N ($N > 1000$ persons) in health state X he would be willing to trade one year of life extension of 1000 healthy individuals for the extension of life by one year for this group. In the second variant (PTO2), the respondent is asked to estimate for how many individuals in health state X he would be prepared to surrender one year of extended life for 1000 individuals in perfect health in exchange for the complete recovery followed by one year of perfect health for the group in the given health state. Hence Murray also incorporates a trade-off element into the operationalization of his PTO protocol.

Murray had a number of indicator conditions assessed during a panel session according to both PTO1 and PTO2. The group process, complete with a discussion of the arguments of the mutual panel members was deemed to be essential to the evaluation to enable the individual panel members to arrive at well-considered assessments, although consensus was not the objective. The reason for subjecting all conditions to both PTO1 and PTO2

assessments was to obtain well-considered valuations by encouraging the participants to deliberate on the reasoning behind one another's assessments.

1.2.2 Respondents in the GBD study: medical experts

The respondents in the original GBD study were health workers from all over the world. It is the opinion of the authors of the GBD that a knowledge of the condition to be evaluated is in fact the factor causing the differences in valuations seen among different groups of respondents.(Murray, 1996b, p.30) A distinction can therefore be made between respondents without a knowledge of the condition to be assessed (such as many members of the general population tend to be) and those with a knowledge of the condition to be valued. This latter group may further be broken down into those who have experienced the condition themselves or are still living in this state (ex patients and patients); persons with an experience of the condition with someone near them [family members and friends of (ex) patients]; and those who have gained a knowledge of the health state through their work (professional health care providers). Various empirical studies have shown that patients and ex patients adapt to their own health state and value this as less severe than non patients. Furthermore, the knowledge of (ex) patients and their family and friends extends to only a limited number of health states.

The GBD investigators ultimately decided to base the weights on the assessments of health care professionals. A knowledge of and insight into the sequelae of the largest possible selection of the conditions to be valued was deemed to be essential. 'Non health care providers could be used but much more time would be required to educate them about each condition.'(Murray, 1996b, p. 37) Opting for doctors also meant that the naturalistic descriptions of the indicator conditions to be assessed in the GBD study could be used.

1.2.3 The valuation procedure in the GBD study: choice of indicator conditions

The first step in the procedure applied in the GBD study was the assessment of 22 indicator conditions by an international panel of health care providers ('Geneva meeting') by means of PTO1 and PTO2. In the term 'indicator conditions' the emphasis is on 'indicator': each condition can be interpreted to represent a dimension of the effects of disease. In this way, there were three conditions for pain: 'severe sore throat' for slight pain, 'angina' for moderate pain, and 'severe migraine' for severe pain. 'Radius fracture in a stiff cast' stood for the functional loss of one arm, 'below the knee amputation' for the loss of one leg, 'paraplegia' for the functional loss of two legs and 'quadriplegia' for the functional loss of all four limbs. The social function was rendered exclusively in a single condition, namely vitiligo on the face, while a predominant social element was captured by 'recto-vaginal fistula'. The indicator conditions also included various neuro-psychiatric disorders: 'mild mental retardation', Down's syndrome', 'unipolar major depression', 'active psychosis', 'dementia'. Other dimensions incorporated into the indicator conditions were: sensory ('blindness', 'deafness'); sexual ('erectile dysfunction') and reproduction ('infertility').

Based on the weights obtained via PTO during the panel session, the spectrum of disabilities was arbitrarily divided into seven more or less homogenous classes (*see table 1.1).*

The weights for hundreds of other conditions were then derived via the far more simple procedure of having the respondents assign them to the appropriate classes. Here, again, deliberation and reconsideration of an estimation once given after hearing the arguments of the other participants was an essential element.

Table 1.1 – Disability classes en severity weights for indicator conditions from the GBD (source: Murray, 1996a)

Disability class	Severity weights	Indicator condition
1	0.00-0.02	Vitiligo on face, weight-for-height less than 2 standard deviations
2	0.02-0.12	Watery diarrhoea, severe sore throat, severe anaemia
3	0.12-0.24	Radius fracture in a stiff cast, infertility, erectile dysfunction, rheumatoid arthritis, angina
4	0.24-0.36	Below-the-knee amputation, deafness
5	0.36-0.50	Rectovaginal fistula, mild mental retardation, Down's syndrome
6	0.50-0.70	Unipolar major depression, blindness, paraplegia
7	0.70-1.00	Active psychosis, dementia, severe migraine, quadriplegia

1.3 Objectives of the Dutch disability weights project

In section 1.1, the reasons behind the need for a coherent set of disease-specific weights were discussed. At the start of the Dutch disability weights study, there were various reasons why it seemed important to align this study with the Global Burden of Disease project. In the first place, the GBD project is impressive due to the well-considered manner in which a coherent set of weights was elicited. As a result, it became possible to combine a huge quantity of data on morbidity and mortality for specific diseases such that the results were able to be compared mutually. The standardized methodology of the GBD project dovetails with the central problem of the Dutch study. In the second place, taking advantage of existing methods offers a basis for comparison and possibilities of refining and validating the methods used. In the third place, the GBD is an authoritative project which may be expected to be copied in (inter)national studies.

The GBD weights were derived, however, for use at the global scale, which means that relatively much attention is spent on conditions with little relevance to the Dutch population's health, such as tropical diseases and malnutrition. Hence little attention is reserved for chronic affluence-related diseases. Moreover, the GBD study was the only one of its kind, which implies that at that time very little was known about the reliability and validity of the methods used. Replication and validation of the GBD study was therefore of essential importance.

The Dutch project on 'Disability Weights for Diseases' was carried out in 1996. The objectives of the project may be summarized as follows:

- to investigate to what extent the method introduced by the GBD study for determining disability weights yields reliable, valid and - for the Netherlands - usable results.
- to derive a coherent set of disability weights with the help of panels of experts for a large number of diseases, for various applications. These applications concern composite public health measures based on both life table techniques and on computer simulation models, and economic evaluations in health care.
- for a start, to apply the weights within the scope of the Public Health Status and Forecast 1997 study *(VTV-97)* in making a tentative estimation of the burden of disease in the Netherlands for a number of important diseases.
- if the weights elicited during the project are deemed sufficiently reliable and valid for general application, to make these weights available to public health research and health services research and to test these in an international forum.

2

DESIGN OF THE STUDY ON DISABILITY WEIGHTS FOR DISEASES IN THE NETHERLANDS

THIS CHAPTER provides a detailed description of the design of the Dutch disability weights study. Weights were derived for a large number of diseases, representing morbidity, or the degree of dysfunction associated with each disease. The set-up is partially derived from that of Murray, as described in *section 1.2*. In the Dutch study, a relatively small number of indicator conditions (in Dutch: *'ijktoestanden'*) were assessed by panels of medical experts, using the PTO valuation method. During the second step, a much larger selection of disease stages were interpolated between the indicator conditions by the individual panel members. At places, it was deliberately chosen to deviate from the method followed by Murray, for example in compiling the list of diseases for which weights were to be derived, in the choice of indicator conditions, as well as in the addition of a standardized description of the health state to be valued and in the application of, next to PTO, a second valuation method, namely the visual analogue scale. The entire study set-up was tested as a pilot study and documented in full.

Prior to the start of the Dutch weights study, the members of the project group participated in November 1995 together with several other Dutch researchers, in a valuation study on behalf of the Global Burden of Disease study led by Murray. This consisted of a panel session in which weights were empirically derived using PTO1 and PTO2 for Murray's 22 indicator conditions (see *section 1.2*).

The following sections explore the design of the study on disability weights performed in the Netherlands.

2.1 Pilot study

There were two points in the protocol established by Murray which primarily merited closer consideration. The first was the person trade-off valuation method (in two variants: PTO1 and PTO2) used by Murray during the panel

session. This method is not customarily applied in the tradition of economic evaluation of medical interventions. More is known about the drawbacks to the more commonly used methods in this connection where conceptual background and practical implementation is concerned [particularly 'standard gamble' (SG), 'time trade-off' (TTO) and 'rating scale' (RS), which latter method includes for example a 'visual analogue scale'(VAS)]. (Krabbe, 1997a) In the second place, Murray had the participants in the panel session assess disease-specific health states, namely naturalistic health states labelled with a diagnostic label rather than generic health states. It may be assumed that the information about the diagnosis assigned to the health state to be assessed will influence the ultimate valuation. (Froberg, 1989; Essink-Bot, 1995) The diagnosis contains implicit information about the prognosis of the condition to be assessed, while the cultural position of a diagnosis (such as AIDS) can also play a role.

In view of these uncertainties in the approach used by Murray, a pilot study was conducted at The Hague on 19 March 1996, with the following objectives:

- to test and compare several methods of valuation (TTO, PTO1, VAS);
- to evaluate the usefulness of the standard EuroQol 5D classification for generating a generic description in a standardized way of the health state corresponding with each disease;
- to gain insight in the difference between presenting health states with and without a diagnostic label.

The test subjects used in the pilot study included the members of the project group and several other researchers active in the valuation of health states. The health states to be assessed were selected based on the availability of data describing the health-related quality of life, applying as criterion the fact that they were important or common diseases in the Netherlands. The health states were assigned standardized descriptions with the help of the standard EuroQol 5D classification system (*see section 3.4.3*), based on the descriptive (functional) health-state data available. (EuroQol Group, 1990; Brooks, 1996) The most important conclusions yielded by this plenary pilot study were:

- the TTO valuation method seemed to offer no practical advantages over PTO.
- PTO1 without PTO2 was not really feasible.
- evaluating health state descriptions with and without a diagnostic label in a single session is confusing.
- the diagnostic label plays a key role in the valuation process; the EuroQol 5D descriptions were perceived by this panel as an addition.
- a dimension for cognitive functioning in the standardized health status descriptions was felt to be essential.

The plenary pilot study was followed up by a written interpolation round. The protocol designed proved on the whole to function well.

2.2 The list of diseases and disease stages

2.2.1 The diseases

The choice of diseases used to elicit the weights was based primarily on the Public Health Status and Forecast 1997 project *(VTV-97)*. In this, 52 diagnostic groups were selected based on their importance to public health in terms of mortality, morbidity and costs. These diagnostic groups cover some 70% of all causes of death, approximately 45-50% of the morbidity in the general medical practice and some 65% of the total health care costs. Apart from the fact that, as a result, not all diseases were included, part of the ill health in the population cannot as such be coupled to a specific disease or condition. This has been partially compensated through the addition of an item to the list, namely that of ADL limitations in the elderly.

2.2.2 Disease stages

The goal of the disability weights project was to elicit weights for the functional sequelae of a number of diseases. The diagnostic groups in the Public Health Status and Forecast 1997 study were described as ICD categories. As such, they are not easily assigned a weight, because an ICD category in terms of (the sequelae of a) disease is often not a homogenous category. 'Dementia' for example causes a broad range of disabilities. It was not deemed possible to ask the respondents to assign a single weight to a condition in its entirety. It would then have to be assumed that each participant a) would be familiar with the entire spectrum of sequelae which can occur with a condition; b) knows the contribution of each disability to the total morbidity burden as a consequence of the condition, which assumes insight into incidence, prevalence and length of illness; and c) is capable of arriving at a single weight by means of an averaging routine for that disorder. It must then also be presumed that all the respondents have an equal command of all this information. This is not a plausible assumption.

It was therefore decided, where necessary and possible, to divide each diagnostic group on the Public Health Status and Forecast list into stages or severity levels. The word *stage* in this connection is to be understood as a more or less homogenous (according to health state, treatment and prognosis) phase, measured in time, of the process of disease. This is therefore a different interpretation than the customary 'stage of disease at diagnosis' used in some clinical contexts. No distinction is made in the following between stage and severity; both are referred to as stage.

The classification into disease stages proved to be a matter of custom tailoring. Diagnostic groups such as malignant growths present successive and irreversible stages in the course of the disease, in principle in each individual, although it should be noted that not every individual passes through each possible stage. There are also diseases which exhibit variations in their course between individuals. Different levels of severity were distinguished for this type of disease. Gastroenteritis, for example, was

divided into a form showing an uncomplicated course and one presenting a complicated course. A comparable kind of heterogeneity between persons with the same disease is seen for example in acute hepatitis B (no symptoms in 50% of the cases, flu-like symptoms in 48% and acute liver failure in 2% of the cases) and in the condition following premature birth (95% have no residual symptoms, 5% are significantly worse off). In cases like these, the symptomatic form was the form assessed (hence symptomatic acute hepatitis B and permanent impairments following premature birth, respectively). The further calculations were adjusted for the sake of heterogeneity.

In the presentation to the panel members of the disease stage to be valued, the complete staging of the disease was made available to them.

2.2.3 Addition of standardized health state description to disease stage

It was assumed that the weighting of health states only described as a diagnostic label would yield less valid results. In the first place, it is highly improbable that a medical expert has real insight into the consequences of all 52 diseases on the list. Even if such medical experts exist, it is strongly unlikely that all the members of the panel would have a comparable level of expertise. Moreover, information about the diagnosis transmits implicit information about the prognosis of the condition to be assessed. These implicit features of a diagnosis are not likely to be known to an equal extent by all the participants. To summarize, there were more than sufficient reasons to standardize the stimulus, i.e. the health state to be valued, by attaching, in addition to the diagnostic label, a description of the appropriate health state. To arrive at this health state description, a six-dimensional extended variant of the EuroQol 5D classification system was used *(see table 3.8)*. The standard EuroQol 5D classification enables generic, instead of disease-specific, descriptions of a health state to be made from five aspects: mobility, self care, usual activities, pain/discomfort, anxiety/depression. Each of these dimensions is divided into three levels: no problems, moderate problems and severe problems. The sixth dimension (cognition) was added to the five standard dimensions for the purpose of this project. The disabilities following from diseases such as dementia, mental retardation and schizophrenia cannot be validly represented by the mere consequences - if any - of cognitive dysfunction on the standard EuroQol 5D items. Clinical experts declared that patients are much more affected by derealization, loss of ‘self’, the feeling of being caged, not being able to think properly, rather than by the consequences of advanced memory loss for e.g. day-to-day activities. The six-dimensional extended EuroQol classification is further referred to as ‘EuroQol 5D+’ A pilot study was conducted on the effect of adding information about cognition to the assessment of generic health states. (Krabbe, 1997b)

2.2.4 Duration

When assessing health states, it remains essential to define the duration of the state to be assessed. Following the GBD methodology, a duration of one year is assumed in the PTO1 for all health states. (Murray, 1996b) This choice is moreover related to the use of one-year prevalence figures in calculating the 'Years Lived with Disability'.* (Ruwaard, 1997, pp.46-51) The health state is considered to remain constant throughout that year. The year is followed in all situations by certain death. Other choices could have been made within the context of PTO.

The assumed duration of a year is a realistic one for the majority of (chronic) disease stages, in that this state can last at least a year. For a number of other diseases, however, a stationary duration of one year is absurd. Assessing one year of influenza or an asthma attack of a year would have yielded bizarre results. Problems are in fact seen with two types of diseases, namely those which occur in an episodic pattern (e.g. asthma, epilepsy, migraine) and conditions with only a brief duration and followed in a majority of cases by a full recovery (e.g. common colds, influenza, gastroenteritis) or by death (septicaemia).

The attacks themselves were not assessed in the episodic group of diseases. Such diseases were described as chronic. Measures to prevent attacks, the side-effects of such measures and the fear of suffering an attack were included in both the description and the EuroQol 5D+ classification (for example: 'severe asthma, i.e. not symptom-free despite maintenance medication'). Brief (infectious) conditions followed by a full recovery were presented for valuation as an *annual profile*, e.g. one year in good health with two weeks of influenza during that year. Hence the entire year, and not simply the episode of illness was presented for assessment. In the annual profiles, the state during the two weeks of influenza was characterized in the EuroQol 5D+ description.

2.2.5 The final lists of disease stages

The classification of the 52 diagnostic groups into homogenous stages, with each stage assigned a representative description of the corresponding health state in EuroQol 5D+ terms, was arrived at as follows. An extra question was inserted into the Public Health Status and Forecast 1997 survey, in which an expert was consulted for information about each diagnostic group. In this extra question, the expert was asked to divide the relevant diagnostic group into a limited number of clinical stages characterized by homogeneity in respect of health state, treatment and prognosis, and to provide a description of each health state in EuroQol 5D+ terms. For most of the diagnostic groups, an additional expert was consulted on specifically the division into stages for the purpose of the disability weights study. Finally, the researchers

* Years Lived with Disability are understood to refer to: the number of life-years lived with a condition, weighted for the severity of this condition, also to be understood as 'severity weighted prevalences'.

themselves drafted a classification into disease stages using EuroQol 5D+ descriptions according to their own insights and on the basis of a limited literature study. Pursuant to the information from these three sources, a tentative classification into stages plus EuroQol 5D+ description was proposed for each diagnostic group. These proposals were evaluated by the members of the project group (horizontal consistency, or plausibility of the stage-classification per disease). After the classifications per disease were adjusted on the basis of this assessment, the entire list of 52 diagnostic groups, divided into disease stages and provided per disease stage with a description in EuroQol 5D+ terms, was valued for plausibility *between* the diseases (vertical consistency) by three independent expert medical generalists. On the basis of these assessments, the list subsequently underwent a final adjustment round. The descriptions of all disease stages with the corresponding EuroQol 5D+ descriptions for all 52 diagnostic groups were listed in an appendix (see Appendix A) (the EuroQol 5D+ coding is explained in *section 3.4.3 and table 3.8*).

Virtually no account was taken of the consequences of more than one disease occurring in a single person (comorbidity) when assigning the weights. A distinction is commonly made between independent and dependent comorbidity. Independent comorbidity is seen when a chance combination of two or more diseases (e.g. arthrosis and heart attack) occurs. A combination can also occur more frequently than expected by mere chance (e.g. diabetes and cardiovascular disease), which is known as dependent comorbidity. Dependent comorbidity arises if one disease forms a risk factor for the other disease, or if the two diseases share a common risk factor. The distinction between independent and dependent comorbidity is probably of no importance to the sequelae in terms of physical, mental and social functioning. Reality is probably more complex than the general assumption that the total level of disability caused by a combination of diseases equals the sum of the disabilities caused by each of the components of the combination. (Verbrugge, 1989)

In the present study, explicit disability weights were elicited for the consequences of several dependent co-morbidities. These included Down's syndrome with other congenital defects (unspecified), diabetes mellitus with neuropathy, and diabetes mellitus with nephropathy. Explicit disability weights for the combined sequelae of a larger number of dependent and independent forms of co-morbidity should be elicited in a follow-up study.

2.3 Selection of the panel members

In conformity with the reasoning of the GBD investigators (*see section 1.2.2.*), it was decided to opt in favour of having physicians participate in the assessment process. When recruiting the participants, efforts were made to select 45 doctors with a broad, general, practical knowledge of medicine. A sufficient ability to reason in the abstract was also needed for performing the task of valuation. No formal method of recruitment was applied. Groups and

organizations approached in mid 1996 included the staff at the Institute for General Medicine *(Instituut voor Huisartsgeneeskunde)* at the Academic Medical Centre (AMC) in Amsterdam, medical specialists at the AMC, the AMC general practitioner training supervisors, the Inspectorate for Health Care (*Inspectie voor de Gezondheidszorg)*, the Health Insurance Funds Council *(Ziekenfondsraad)*, the Dutch College of General Practitioners *(Nederlands Huisartsen Genootschap)* and interviewers from the assessment of the euthanasia reporting procedure.

The participants took part in a plenary panel session in three groups. Using the PTO valuation method, 16 disease stages (*indicator conditions*) were assessed in each panel session.

It was separately investigated whether the background of the participants was of influence in deriving the disability weights. To this end, in addition to the three panels of physicians with a broad medical knowledge and experience, a 'lay' panel was put together, made up of participants boasting an academic background but no medical knowledge. A total of 15 members of staff (primarily economists, sociologists, political scientists and lawyers) on the Scientific Council for Government Policy (*Wetenschappelijke Raad voor het Regeringsbeleid)* were asked in March 1997 whether they were interested in participating in this study. The structure of the lay panel session was identical to that of the panels of medical experts.

Person Trade-Off - PTO1

In the first variant of the Person Trade-Off method, you are asked to undertake a thought experiment in which you trade off life years of healthy people for life years of individuals who are not in perfect health.

Imagine the following:

You are a decision maker. You have exactly enough funds for a single health intervention. You have a choice between two mutually exclusive health interventions.

If you opt for intervention A, the life of 1,000 individuals will be extended by exactly one year. After that year they will all die. If you do not choose this intervention, these people will all die immediately.

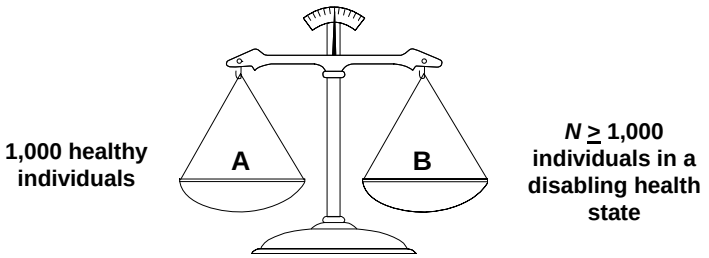
Alternatively, your scarce funds may be used to purchase health intervention B. Opting for B means that the life of *N* individuals in the less than perfect health state X would be extended by exactly one year. After that year, they will all die. Not choosing intervention B means that the persons in health state X will all die immediately.

Example:

The choice is in the first instance between one year of life extension for 1,000 healthy individuals (intervention A) and one year of life extension of 2,000 blind people (intervention B). If you opt for B, you will be faced with a new choice in which the number of blind individuals whose life can be extended with intervention B is reduced to, e.g. 1,500. If you decide to purchase A, the number of blind individuals will be raised. This process of choosing is continued until you are no longer able to make a choice between the two interventions: your indifference point.

In summary:

PERSON TRADE-OFF 1: the number (*N*) of individuals in health state X for whom one year of life extension is equal in your eyes to one year life extension for 1,000 healthy individuals. The number *N* is always bigger or equal to 1,000.



A PTO1 of 1,000 implies that you value the given health state A as equal to 'perfect health'. A PTO1 of 1,000,000 (1 million) means that you value the given health state X as extremely bad. Your PTO1 valuations may be anywhere between these two extremes.

Figure 2.1 – PTO1 Instruction

Person Trade-Off - PTO2

In the second variant of the PERSON TRADE-OFF METHOD (PTO2) you are again invited to participate in a thought experiment. This time, you are asked to make a trade off between life extension for healthy individuals and an improvement in the quality of life of individuals in a disabling health state.

Imagine the following:

You are a decision maker. You have exactly enough funds for a single health intervention. You have a choice between two mutually exclusive health interventions.

If you opt for intervention A, the life of 1,000 individuals will be extended by exactly one year. After that year they will all die. If you do not choose this intervention, these people will all die immediately.

Alternatively, your scarce funds may be used to purchase health intervention B. With intervention B, *N* individuals in health state X will undergo a complete recovery. Intervention B will allow them to live for one year in perfect health. After that year, they will all die. If you choose not to purchase intervention B, they will live for one year in health state X, after which they will all die.

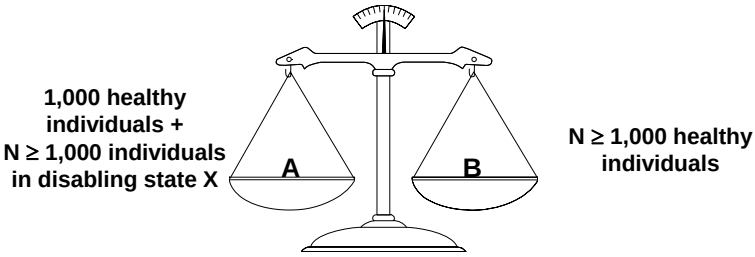
A decision maker purchasing intervention B trades off 1,000 healthy life years for the full recovery of *N* individuals in health state X.

Example:

The choice is in the first instance between one year of life extension for 1,000 healthy individuals (intervention A) and the full recovery of 2,000 blind people (intervention B). If you opt for B, you will be faced with a new choice in which the number of blind individuals able to regain perfect health with intervention B is reduced to, e.g. 1,500. If you decide to purchase A, the number of blind individuals who regain their sight will be raised. This process of choosing is continued until you are no longer able to make a choice between the two interventions: your indifference point.

In summary:

PERSON TRADE-OFF 2: the number (*N*) of individuals in health state X for whom a complete recovery, followed by one year of perfect health is equal in your eyes to one year life extension for 1,000 healthy individuals. The number *N* is always bigger or equal to 1,000.



A PTO2 of 1,000 implies that you value the given health state X as equal to ‘perfect health’. A PTO2 of 1,000,000 (1 million) means that you value the given health state X as extremely bad. Your PTO2 valuations may be anywhere between these two extremes.

Figure 2.2 – PTO2 Instruction

Table 2.1 – Conversion table PT01–PT02

PTO1 →	PTO2	PTO1 →	PTO2	PTO1 →	PTO2
1.001	1.001.000	1.350	3.857	5.500	1.222
1.002	501.000	1.400	3.500	6.000	1.200
1.003	334.333	1.450	3.222	6.500	1.182
1.004	251.000	1.500	3.000	7.000	1.167
1.005	201.000	1.550	2.818	7.500	1.154
1.006	167.667	1.600	2.667	8.000	1.143
1.007	143.857	1.650	2.538	8.500	1.133
1.008	126.000	1.700	2.429	9.000	1.125
1.009	112.111	1.750	2.333	9.500	1.118
1.010	101.000	1.800	2.250	10.000	1.111
1.011	91.909	1.850	2.176	11.000	1.100
1.012	84.333	1.900	2.111	12.000	1.091
1.013	77.923	1.950	2.053	13.000	1.083
1.014	72.429	2.000	2.000	14.000	1.077
1.015	67.667	2.050	1.952	15.000	1.071
1.016	63.500	2.100	1.909	16.000	1.067
1.017	59.824	2.150	1.870	17.000	1.063
1.018	56.556	2.200	1.833	18.000	1.059
1.019	53.632	2.250	1.800	19.000	1.056
1.020	51.000	2.300	1.769	20.000	1.053
1.021	48.619	2.350	1.741	21.000	1.050
1.022	46.455	2.400	1.714	22.000	1.048
1.023	44.478	2.450	1.690	23.000	1.045
1.024	42.667	2.500	1.667	24.000	1.043
1.025	41.000	2.600	1.625	25.000	1.042
1.030	34.333	2.700	1.588	30.000	1.034
1.040	26.000	2.800	1.556	35.000	1.029
1.050	21.000	2.900	1.526	40.000	1.026
1.060	17.667	3.000	1.500	45.000	1.023
1.070	15.286	3.100	1.476	50.000	1.020
1.080	13.500	3.200	1.455	55.000	1.019
1.090	12.111	3.300	1.435	60.000	1.017
1.100	11.000	3.400	1.417	65.000	1.016
1.110	10.091	3.500	1.400	70.000	1.014
1.120	9.333	3.600	1.385	75.000	1.014
1.130	8.692	3.700	1.370	80.000	1.013
1.140	8.143	3.800	1.357	85.000	1.012
1.150	7.667	3.900	1.345	90.000	1.011
1.160	7.250	4.000	1.333	95.000	1.011
1.170	6.882	4.100	1.323	100.000	1.010
1.180	6.556	4.200	1.313	150.000	1.007
1.190	6.263	4.300	1.303	200.000	1.005
1.200	6.000	4.400	1.294	250.000	1.004
1.210	5.762	4.500	1.286	300.000	1.003
1.220	5.545	4.600	1.278	350.000	1.003
1.230	5.348	4.700	1.270	400.000	1.003
1.240	5.167	4.800	1.263	450.000	1.002
1.250	5.000	4.900	1.256	500.000	1.002
1.300	4.333	5.000	1.250	1.000.000	1.001
PTO1 →	PTO2	PTO1 →	PTO2	PTO1 →	PTO2

Visual Analogue Scale (VAS)

On the VISUAL ANALOGUE SCALE, you are asked to give a direct rating of health states. You will receive 16 cards. A health status and a letter are written on each card. The standardized description of the health state in question is printed on the back of each card.

Valuations using VAS are performed in two steps.
1) During the first step, you are asked to rank these health states, from the health state 'most highly valued' to the 'least valued' health state. Lay your cards down in that order.
2) You are subsequently asked to assign a score to the health states which you had previously ordered. To this end, a thermometer (VAS) is provided marked from 0 to 100. 0 is the worst possible rating ('dead') and 100 is the best possible valuation ('healthy'). Your VAS valuations are to be somewhere in between these two extremes.

Each card has a description of the health state which is coded with a letter. You can assign each health state a value on the thermometer by placing an arrow before the number in question on the scale (valuation) with next to this the corresponding letter of the health state.

Example:

Choose the health state 'blindness'.
In step 2, you must rate the severity of 'blindness' as a health state somewhere between 'healthy' and 'dead'. In the example shown, blindness was fictitiously valued at 73.
Some four other valuations of health states (coded A, C, G, and K) are also indicated in the example.

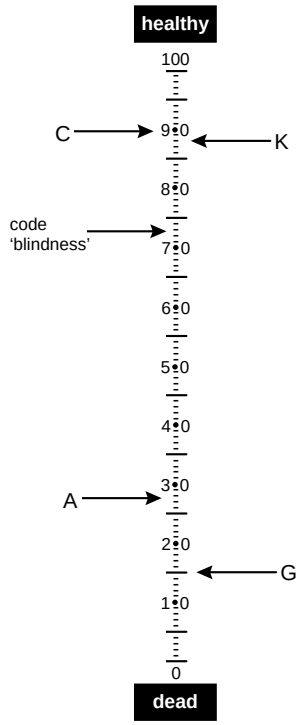


Figure 2.3 – Instruction VAS

2.4 Valuation methods: PTO, VAS

Just as in the GBD study, use was made in the panel sessions of PTO1 and PTO2. The exact operationalization is shown in *figure 2.1 and figure 2.2*.

PTO1 and PTO2 may under certain assumptions be converted to one another in a mathematical equation:

$$PTO2 = 1000 / (1 - \frac{1000}{PTO1})$$

The panel members were given a conversion table for converting PTO1 to PTO2 and vice versa (*see table 2.1*).

The visual analogue scale was applied as the third valuation method during the panel sessions (*see figure 2.3*). PTO is a relatively new method. Comparison with the more commonly used VAS method can enable the knowledge of the properties of PTO to be expanded. The VAS is a kind of thermometer running from 0 (*dead*) to 100 (*healthy*). Participants were asked to rate each condition by marking its place in his or her opinion on the VAS.

2.5 Selection of the indicator conditions

The indicator conditions assessed in the panel sessions yield the values to calibrate what is further referred to as the disability scale. The indicator conditions were carefully selected on the basis of three criteria:

- the indicator conditions were to comprise the most important diagnoses from the Public Health Status and Forecast list, in order to at least rank these as well as possible on the disability scale.
- the indicator conditions were to be well-known diseases, so that participants could be expected to have a clear and comparative idea of the condition to be assessed.
- the indicator conditions were collectively to encompass the entire spectrum from near death to perfect health.

The diseases to yield the indicator conditions were chosen from the Public Health Status and Forecast 1997 list of 52 diagnostic groups on the grounds of the first two criteria. In choosing the disease stages to be valued, the statistical model of Van Busschbach et al. was used for the purpose of assessing the third criterion. (Busschbach, 1997) With the help of this model, valuations can in principle be estimated for all the diseases stages able to be described by the (standard 5D) EuroQol system. The model can thus offer an impression of the expected distribution of the indicator conditions on the scale. In the end, some 16 indicator conditions were selected (*see figure 2.4*).

Figure 2.4 – Description of the indicator conditions

Indicator condition	EuroQol 5D+ description
VISION DISORDER Patients with severe vision disorders, i.e. unable to read small newspaper print, great difficulty or unable to recognize faces at 4m. distance.	↗ No problems in walking about ↗ Some problems with washing and dressing self ↗ Unable to perform usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No pain or discomfort ↗ Moderately anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
DIABETES MELLITUS Patiens with uncomplicated diabetes mellitus	↗ No problems in walking about ↗ No problems with washing and dressing self ↗ No (90%) or some (10%) problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No (90%) or some (10%) pain or discomfort *(in this case particularly discomfort, no pain)* ↗ Not (90%) or moderately (10%) anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
LOW BACK PAIN Patients with low back pain	↗ Some problems in walking about ↗ No problems with washing and dressing self ↗ Some problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ Moderate pain or discomfort *(in this case particularly pain, no discomfort)* ↗ Not anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
SCHIZOPHRENIA Patients with schizophrenia, several psychotic episodes, severe and increasing permanent impairments.	↗ Some problems in walking about ↗ Unable to wash or dress self ↗ Unable to perform usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ Extreme pain or discomfort *(in this case particularly discomfort, no pain)* ↗ Extremely anxious or depressed ↗ Extreme problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ) (in this case not the IQ-level)*

CORONARY HEART DISEASE Patients with mild stable angina pectoris (NYHA 1 - 2)	↗ No problems in walking about ↗ No problems with washing and dressing self ↗ No problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ Moderate pain or discomfort *(in this case particularly pain, no discomfort)* ↗ Not anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
DEMENTIA Patients with severe dementia (permanent supervision required)	↗ Some problems in walking about (50%) of confined to bed (50%) ↗ Unable to wash or dress self ↗ Unable to perform usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No pain or discomfort ↗ Moderately (50%) of extremely (50%) anxious or depressed ↗ Extreme problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
DENTAL DISEASE Patients with periodontal disease (gingivitis)	↗ No problems in walking about ↗ No problems with washing and dressing self ↗ No problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No pain or discomfort ↗ Not anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
ACCIDENTS AND INJURIES Patients with paraplegia, in permanent stage	↗ Some problems in walking about (wheelchair, 80%) or confined to bed (20%) ↗ Some problems (80%) of unable (20%) wash or dress self ↗ Some problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No (80%) or moderate (20%) pain or discomfort *(in this case particularly discomfort, no pain)* ↗ Not (80%) or moderately (20%) anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*

STROKE (CVA) Patients after stroke, moderate permanent impairments	↗ Some problems in walking about ↗ Some problems with washing and dressing self ↗ Some problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ Moderate pain or discomfort ↗ Moderately anxious or depressed ↗ Some problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ) (in this case particularly memory and concentration)*
DEPRESSION Patients with a mild depression, i.e. some limitations in work and social functioning	↗ No problems in walking about ↗ No problems with washing and dressing self ↗ Some problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No pain or discomfort ↗ Moderately anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
COLORECTAL CANCER Patients with colorectal cancer, irradically removed or disseminated	↗ No (80%) or some (20%) problems in walking about ↗ No (80%) or some (20%) problems with washing and dressing self ↗ Some problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ Moderate (80%) or extreme (20%) pain or discomfort ↗ Extremely anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
ASTHMA/COPD Patients with mild to moderate asthma, i.e. symptom-free with or without maintenance medication	↗ No problems in walking about ↗ No problems with washing and dressing self ↗ No (75%) or some (25%) problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No pain or discomfort ↗ Not (95%) or moderately (5%) anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*

RHEUMATOID ARTHRITIS Patients met severe reumatoide arthritis	↗ Some problems in walking about (80%) or confined to bed (20%) ↗ Some problems (80%) or unable to (20%) wash and dress self ↗ Some problems (80%) or unable to (20%) perform usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ Extreme pain or discomfort *(in this case particularly pain, no discomfort)* ↗ Extremely anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
BREAST CANCER Patients with breast cancer, clinically disease-free after the first year	↗ No problems in walking about ↗ No problems with washing and dressing self ↗ No problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ Moderate pain or discomfort ↗ Moderately anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
HEALTH PROBLEMS IN MATURELY BORN CHILDREN Children with permanent impairments after asphyxia at birth (APGAR < 7 after 5 minutes)	This leads to the follow permanent stage in 10% of the patients (90% no disadvantageous effects): ↗ Some problems in walking about ↗ Some problems with washing or dressing self ↗ Some problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No pain or discomfort ↗ Moderately anxious or depressed ↗ Some problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*
ADL-LIMITATIONS Elderly with moderate to severe ADL-limitations	↗ Some problems in walking about ↗ Some problems with washing or dressing self ↗ Some problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)* ↗ No pain or discomfort ↗ Not anxious or depressed ↗ No problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*

2.6 Description panel session

The panel sessions (three in total) were held on 9, 16 and 25 October 1996 in the Academic Medical Centre (AMC) in Amsterdam. Fifteen participants were invited to each panel session. Prior to these sessions, the protocol for the sessions was tested on several members of staff in the department for Clinical Epidemiology and Biostatistics at the AMC.

Prof. Dr. L.J. Gunning-Schepers acted as panel leader in all three sessions; the researchers in charge of the study (MS, MLEB) attended each session as observers. At the start of each panel session, the panel leader explained the backgrounds and objectives of the study. Each panel member was then introduced to the procedure individually by going through an assessment of the condition 'severe vision disorder' using PTO1 followed by PTO2. A ping-pong procedure was used in both PTO assignments, as a means to determine the indifference point of the panel member. Panel members were then invited to air their thoughts and to explain the reasoning behind a particular choice. Next, PTO1 and PTO2 were individually adjusted for consistency. After this initial round of individual practice with PTO1 and PTO2, which took up a large portion of the morning session, the subsequent indicator conditions were valued by all participants simultaneously. Each wrote his assessment on a white board, after which the individual assessments were compared by the panel members. The arguments behind the valuations were deliberated on in a discussion forum led by the panel leader. The object of these discussions was not (as was the case with Murray) to achieve group consensus, but to enable each participant to come to a well-considered, well-argued valuation. The valuations could be revised after the discussion, an option which was in many cases utilized. The first and the second, revised valuation were both written on a PTO form (*see figure 2.5*). On the left-hand side of the form, the diagnostic group was broken down into disease stages (disease model); below this, the condition to be assessed was printed in bold print in a separate box, under which the EuroQol 5D+ description was provided. The right-hand side of the form was reserved for filling in the PTO values. Assessing the 16 indicator conditions with PTO1 and PTO2 took up the entire morning and part of the afternoon.

During all three panel sessions, panel members asked questions about the prognosis, adaptation, the reference group, the context and duration. The instruction to the panel members was that a prognosis of a disease stage was not permitted to be included in the weighting process, but the uncertainty of the patient about the prognosis could. This uncertainty had also been, where possible, factored into the EuroQol 5D+ descriptions. Relating to adaptation, people are capable of adapting to changed living conditions. The adaptation of people to a life in a (chronic) less than optimum health state was, however, not to be valued. (Murray, 1996b) The reference group of healthy persons was always in all respects comparable to the group of persons in a poorer state of health, except, naturally, for the disease stage to be valued. The context in which the assessment were made was the situation in the Netherlands, with all the attendant facilities available. The health status was

therefore given the treatment standard in the Netherlands, hence in the case of a vision disorder, adequate correction (glasses) was assumed. The duration of the health state to be valued was in all cases one year.

After the PTO valuations had been performed, the 16 indicator conditions were ranked according to a new procedure and valued on a 'visual analogue scale' (VAS). To that end, the panel members were handed the indicator conditions on cards (containing both diagnosis and disease stage plus description of the health state). This procedure was not all that time-consuming as the discussion on disease stages had already been held.

Finally, the panel members were asked to write down the rankings of the PTO and the VAS valuations on a form and to reconcile these for the purpose of consistency where necessary. Each individual could adjust both the PTO valuation and that of the VAS. During this phase, panel members once again had the opportunity to revise their earlier assessments.

2.7 Interpolation

Based on the average PTO valuations of all the respondents, the indicator conditions were ranked on a disability scale (*see figure 3.2*). During a written follow-up round, each of the participants was subsequently asked to interpolate 30 new disease stages on this disability scale. Of these, six were the same for all respondents, being: terminal disease stage, severe heart failure, multiple sclerosis in relapsing-remitting phase, severe hearing disorders in the elderly, influenza in annual profile, light to moderate post traumatic stress disorder. These six disease stages formed the common core presented to all 38 panel members in the interpolation procedure. The 'common core' was selected according to the following criteria: the spread of the disease stages in the common core on the disability scale, 'difficult' disease stages (example of a diagnosis in annual profile; not further specified terminal disease) and a poor overlap with the indicator conditions. The other 153 disease stages to be interpolated were distributed among the participants. All the participants were given a different set of some 30 disease stages to be assessed, following a factorial design. All disease stages were interpolated by a total of six panel members.

VISION DISORDER

Vision disorder is subdivided in the following three stages:

(1) mild vision disorder,
(2) moderate vision disorder,
(3) severe vision disorder.

We now ask you to value the following health state:

Patients with severe vision disorder, i.e. unable to read small newspaper print, great difficulty or unable to recognize faces at 4m. distance

Description

- No problems in walking about
- Some problems in washing or dressing self
- Unable to perform usual activities
  (eg. work, study, housework, family or leisure activities)
- No pain or discomfort
- Moderately anxious or depressed
- No cognitive impairment
  (eg. memory, concentration, disorganisation, IQ-level)

PTO1

The number of persons in this health state whose life is extended for exactly one year at which you are indifferent between this option and extending the life of 1000 healthy persons for exactly one year.

1e PTO1

2e revised PTO1

3e definitive PTO1

PTO2

The number of persons in this health state who would be raised to disability-free life for exactly one year at which you are indifferent between this option and extending the life of this number of persons in this health state plus 1000 healthy persons for exactly one year.

1e PTO2

2e revised PTO2

3e definitive PTO2

Figure 2.5 – PTO form

Figure 2.5 – PTO form

2.8 Test-retest

Two months after the interpolation forms were sent to the respondents, the six common core disease stages were submitted in writing for a second assessment. The respondents were asked to interpolate these once again, in order to determine the stability of the interpolations. It was assumed that the respondent would be unable to recall exactly how they had valued the common core disease stages two months previously. The common core disease stages were not recognizable as such for the respondents during the first round.

2.9 Analysis of the data

The analysis of the data of the disability weights study comprised a number of phases. During the panel sessions, the PTO and VAS valuations were derived for the indicator conditions. Weights for the indicator conditions were calculated based on the PTO valuations. The weights for the indicator conditions determined the disability scale, that was subsequently used for the interpolations. The weights for the non indicator disease stages were elicited from the interpolations. These subsequent steps will be explained below, followed by the reliability and validity analyses.

2.9.1 Calculating the disability weights

The individual PTO and VAS assessments from the panel sessions were converted to weights using the following formulas:

$$V_{PTO}(Q) = [1000/PTO1] = [1 - 1000/PTO2]$$

$$V_{VAS}(Q) = [VAS/100]$$

where the weight V in $V_{PTO}(Q)$ is based on the PTO1 or PTO2 valuation for indicator condition Q and $V_{VAS}(Q)$ idem for the VAS valuation.

The average disability weights for the indicator conditions were calculated from the individual weights assigned by all participants in the three panel sessions collectively. The weights elicited from the interpolation session were calculated in a number of steps (*see table 2.2*).

- Step 1: The weights of the indicator conditions were calculated as the average $V_{PTO}(Q)$ (see above);
- Step 2a: An interpolation between two indicator values was assigned as weight the average value of the weights of the indicator conditions in question.

- Step 2b: An interpolation equal to an indicator value yields the weight for that value; the end points of the scale (the best imaginable and worst imaginable health state) could also be used;
- Step 3: The final weight assigned to an interpolation condition is the average of the weights thus obtained from the individual panel members.

Table 2.2 - Calculation of disability weights from the interpolations

step	example	operation
1	the mean disability weight for indicator condition X is $V_{PTO}(Q_X) = 0.70$; and the mean disability weight for indicator condition Y is $V_{PTO}(Q_Y) = 0.50$.	-
2a	evaluator A interpolated disease stage Z between indicator conditions X and Y	disability weight for disease stage Z: $V(Q_Z)_A = V(Q_X;Q_Y) = (0.70 + 0.50) / 2 = 0.60$
2b	evaluator B valued disease stage Z equal to indicator condition X	disability weight for disease stage Z: $V(Q_Z)_B = V(Q_X) = 0.70$
3	-	eventual disability weight for disease stage Z: mean = $[V(Q_Z)_A + V(Q_Z)_B]/2 = 0.65$

2.9.2 Reliability and validity of the weights

The reliability of the weights assigned to the indicator conditions derived from the panel sessions (inter panel reliability) was investigated in two ways:

- as the extent of agreement between the three panels on the average weights (multivariate analysis of variance)
- as the extent of agreement between the three panels on the ranking of the weights (Spearman rank correlation coefficient).

The average weights assigned to the indicator conditions by the three panels were not expected to diverge all that much, while the disability weights assigned the indicator conditions were also expected to be more or less the same for all three panels.

The reliability of the weights of the disease stages in the interpolation procedure were not investigated for all disease stages, as not all disease stage had been interpolated by all panel members. The reliability of the interpolations of the common core disease stages was studied in two ways:

- by computing the extent of agreement between the interpolations of individual panel members with the average of the remainder of the group

(individual-rest correlation, comparable to item-rest correlation in which the panel members are the 'items').
- by the test-retest reliability [comparison of mean disability weights using the Student t-test and the (ranking) correlations between the weights of the test and the retest].

The validity of the weights was studied by:
- comparison with the disability weights in the GBD study.
- considering the ranking order of the disability weights (mutual comparison within and between diagnoses).
- estimating the weights for the disease stages based on the standardized EuroQol 5D+ descriptions with the help of a statistical model.

Generally, no exact criteria for the validity of the disability weights can be provided. A reasonable correlation with the weights from the GBD study would appear essential, with due observance of the differences to be anticipated between weights applying on a worldwide, and those on a national scale. Such differences can be expected to occur as a result of, e.g. different treatment options and adaptations to different cultural situations. This implies that the differences found, if any, must be explainable.

The mutual coherence of the set of weights is a second criterion for validity: mutual comparisons both within and between diagnostic groups is to prove that a ranking of the disease stages on the basis of the disability weights is plausible. For example: 'breast cancer, clinically disease-free after the first year' should, logically speaking, be valued as less severe than 'disseminated breast cancer', but more severe than 'gingivitis'.

A third criterion for validity is that of the comparability between the weights derived by the panels and the weights estimated with the help of a statistical model based on valuations for functional health states according to the EuroQol 5D+ system. Again, a reasonable extent of agreement should be found. The estimated weights and those derived by the panels may be expected to demonstrate at least a similar ranking. However, deviations may be seen in particular for those disease stages in which perceived prognosis plays an important role, for relatively unknown conditions and for conditions yielding severe cognitive problems.

3

RESULTS: DISABILITY WEIGHTS FOR DISEASES

THE RESULTS of the Dutch study on disability weights for diseases are presented in the following five sections. First, some relevant information is provided about the various panel members, who were all doctors. Next, the results of the panel sessions are reported: the disability weights for the indicator conditions and the reliability of these weights. The results of the interpolation session are subsequently given, and the reliability of these results. The validity of the resulting disability weights is then discussed. Finally, the results of the lay panel session are presented and compared with those of the panels of physicians.

3.1 Description of the panels

Some 38 physicians divided among three panels participated voluntarily in the study, 28 men (74%) and 10 women (26%) (*see table 3.1*). The average age was 47.7 (SD=9.2). There were no statistically significant differences between the panels in respect of age and sex.

The average number of years of practical medical experience was 15.5 (SD = 10.1). In total, 21 panel members were still involved in direct patient care, while for the other panel members this was one or more years ago (average 7.4 years ago). Here, again, no differences were found between the panels. The majority of panel members had experience with practical medical work in the field of general medicine (74%) and/or clinical medicine (30%).

In addition to direct patient care 30 of the panel members held positions with no direct bearing on patient care, such as scientific research (53%), medical teaching (55%), and public health care (13%).

All the panel members performed the weight procedure for all 16 indicator conditions. The Person Trade-Off method was accepted by all participants as a valuation method, in some cases after much initial difficulty. Nonetheless, the PTO method (*see section 3.2*) proved not to yield

interpretable results for all panel members. The problems mentioned by the panel members were varied, and included:

- the forced choice with PTO between two groups of individuals, of which only one is granted a life extension of one year (reluctance to 'play God')
- the inability to work with the large numbers yielded by PTO1 or PTO2.
- the unrealistic option that individuals in PTO2 were able to regain perfect health and had no adaptation problems with the healthy status (e.g. in the case of severe vision disorder).
- the exclusion of the prognosis in the assessment process (although individual uncertainty about the prognosis had to be included)
- the exclusion of the adaptation by individuals to their disability in the assessment process ('happy slave' argument).
- the rather simple standardized description of the health states, comprising only three levels, of which the median in particular was considered too broad.
- the lack of representativeness of the standardized description of the health state, as this description did not always correspond with the average patient envisaged by a panel member.

The standardized descriptions of health states had a clear added value in describing the disease stages to be valued, despite the restrictions mentioned. Among other things, this allowed the disease stages presented to be discussed in comparable terms and the dimension most decisive in valuing a health state to be recognized.

Table 3.1 – Background characteristics of the panel members

		panel			total (n = 38)
		I (n=14)	II (n=12)	III (n=12)	
gender [a]	male/	12	9	7	28
	female	2	3	5	10
age [b]	mean	48.6	48.4	46.0	47.7
	SD	9.9	9.3	8.9	9.2
involved in direct patient care [c]	yes/	7	6	8	21
	no	7	6	4	17
practical medical experience (years)[d]	mean	17.1	17.0	12.0	15.5
	SD	11.8	10.2	7.4	10.1

a $\chi^2_{(df=2)} = 2.51$ $p = 0.28$.
b $F_{(df=2.35)} = 0.30$ $p = 0.74$.
c $\chi^2_{(df=2)} = 0.92$ $p = 0.63$.
d $F_{(df=2.35)} = 1.04$ $p = 0.36$.

3.2 Results of the panel sessions

During the panel sessions, the indicator conditions were valued according to two valuation methods: person trade-off (PTO) and visual analogue scaling (VAS). The weights derived based on the PTO assessments for the 16 selected indicator conditions form the scale values used in the interpolation session to elicit disability weights for the other conditions.

3.2.1 Disability weights for the indicator conditions

On inspecting the results, the valuations of four panel members proved unusable due to inacceptable invariance. The assessments of these four panel members were therefore not included in the calculation of the weights for the

Table 3.2 – Disability weights for the indicator conditions: mean, standard deviation and median for person trade-off methods (PTO1/PTO2) and VAS (N = 34)

disease stage	PTO1/PTO2			VAS			
	mean	*(SD)*	median	mean	*(SD)*	median	Student *t* [a]
periodontal disease (gingivitis)	1.00	*(0.00)*	1.00	0.99	*(0.02)*	0.99	3.62**
mild/ moderate asthma	0.97	*(0.06)*	0.98	0.92	*(0.05)*	0.95	4.83**
low back pain	0.94	*(0.07)*	0.97	0.87	*(0.09)*	0.87	5.67**
diabetes mellitus (uncomplicated)	0.93	*(0.14)*	0.96	0.88	*(0.07)*	0.90	2.81**
mild angina pectoris	0.92	*(0.08)*	0.95	0.84	*(0.09)*	0.85	5.63**
moderate ADL-limitations	0.89	*(0.16)*	0.95	0.82	*(0.11)*	0.84	3.63**
mild depression	0.86	*(0.16)*	0.91	0.77	*(0.13)*	0.80	3.48**
breast cancer (clinically disease free)	0.74	*(0.23)*	0.82	0.66	*(0.20)*	0.68	2.01NS
severe vision disorder	0.57	*(0.27)*	0.59	0.62	*(0.17)*	0.66	-1.43NS
state after asphyxia	0.51	*(0.31)*	0.56	0.62	*(0.13)*	0.60	-1.84NS
paraplegia	0.43	*(0.24)*	0.45	0.54	*(0.15)*	0.52	-2.33*
state after stroke, moderate permanent impairments	0.37	*(0.26)*	0.33	0.51	*(0.18)*	0.54	-4.00**
colorectal cancer (disseminated)	0.17	*(0.22)*	0.10	0.33	*(0.17)*	0.34	-4.76**
severe rheumatoid arthritis	0.06	*(0.06)*	0.05	0.30	*(0.15)*	0.28	-9.50**
severe dementia	0.06	*(0.14)*	0.02	0.22	*(0.15)*	0.20	-5.34**
severe schizophrenia	0.02	*(0.03)*	0.01	0.14	*(0.13)*	0.10	-6.40**

[a] Paired t-test with df=32 (VAS values missing for one panel member); ** $p < 0.01$; * $p < 0.05$; NS $p \geq 0.05$.

indicator conditions. This implies that the disability weights were calculated over 34 panel members.

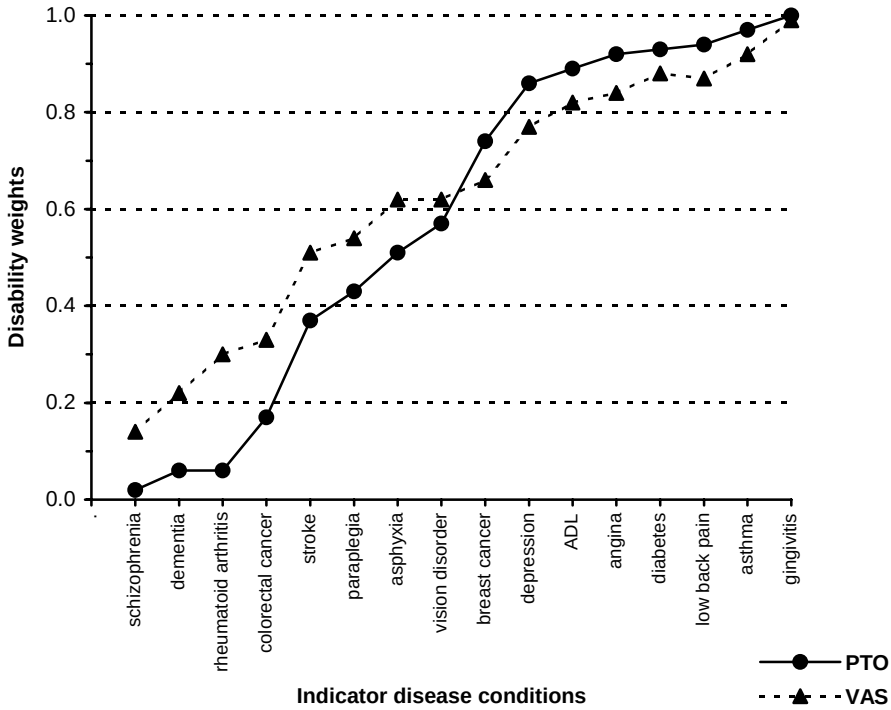


Figure 3.1 – Disability weights for indicator conditions based on average valuations

Pursuant to the instructions, the weights for PTO1 and PTO2 should be identical, as the final PTO1 and PTO2 valuations (*see PTO form; figure 2.5*) have been adjusted for consistency. For this reason, the results for PTO1 and PTO2 are not presented separately. It is worth noting that panel members did not in many cases perceive their assessment task for PTO1 and PTO2 as identical.

The results (*see table 3.2*) reveal that on an aggregated level the rankings of the disability weights are virtually the same for PTO and VAS, both average (Spearman rank correlation rs = .99) and the median (Spearman rank correlation rs = .98). This is hardly surprising, as the ranking of PTO and VAS has been forcibly rendered equivalent at the individual level.

The absolute values of the weights for PTO and VAS do vary: the weights based on the VAS are lower at the top of the scale and higher at the bottom of the scale than the weights based on PTO. Moreover, the difference at the bottom is bigger than at the top (*also see figure 3.1*). A paired t-test of the difference between PTO and VAS disability weights based on the average achieved significance for the majority of values; only for severe vision disorder, diabetes (uncomplicated) and breast cancer (clinically disease-free after the first year) no difference is seen between PTO and VAS weights.

3.2.2 Reliability of the values found for the indicator conditions

Inter-panel reliability of the weights derived for the indicator conditions was investigated in two ways: (a) as the degree of similarity between the three

Table 3.3 – Inter-panel comparison of disability weights

indicator condition	panel									ANOVA[a] F
	I (n=12)			II (n=10)			III (n=12)			
	mean	*(SD)*	rank	mean	*(SD)*	rank	mean	*(SD)*	rank	
periodontal disease (gingivitis)	1.00	*(0.00)*	**1**	1.00	*(0.00)*	**1**	1.00	*(0.00)*	**1**	0.39^{NS}
mild/ moderate asthma	0.93	*(0.09)*	**3**	0.99	*(0.00)*	**2**	0.98	*(0.00)*	**2**	3.49^{*}
low back pain	0.94	*(0.04)*	**2**	0.94	*(0.10)*	**5**	0.93	*(0.06)*	**3**	0.21^{NS}
diabetes mellitus (uncomplicated)	0.89	*(0.20)*	**5**	0.97	*(0.04)*	**3**	0.92	*(0.11)*	**4**	0.72^{NS}
mild angina pectoris	0.93	*(0.09)*	**3**	0.94	*(0.07)*	**5**	0.89	*(0.08)*	**7**	1.11^{NS}
moderate ADL-limitations	0.86	*(0.19)*	**6**	0.91	*(0.15)*	**7**	0.90	*(0.14)*	**5**	0.29^{NS}
mild depression	0.74	*(0.22)*	**7**	0.95	*(0.03)*	**4**	0.90	*(0.05)*	**5**	7.50^{**}
breast cancer (clinically disease free)	0.74	*(0.27)*	**7**	0.73	*(0.27)*	**8**	0.76	*(0.18)*	**8**	0.03^{NS}
severe vision disorder	0.55	*(0.33)*	**9**	0.63	*(0.30)*	**9**	0.52	*(0.20)*	**10**	0.45^{NS}
state after asphyxia	0.23	*(0.24)*	**11**	0.61	*(0.28)*	**10**	0.71	*(0.16)*	**9**	14.56^{**}
paraplegia	0.45	*(0.27)*	**10**	0.43	*(0.20)*	**11**	0.41	*(0.26)*	**12**	0.09^{NS}
state after stroke, moderate permanent impairments	0.20	*(0.22)*	**12**	0.38	*(0.24)*	**12**	0.52	*(0.21)*	**11**	6.44^{**}
colorectal cancer (disseminated)	0.11	*(0.11)*	**13**	0.25	*(0.32)*	**13**	0.17	*(0.22)*	**13**	1.00^{NS}
severe rheumatoid arthritis	0.07	*(0.07)*	**14**	0.05	*(0.05)*	**15**	0.06	*(0.05)*	**14**	0.19^{NS}
severe dementia	0.07	*(0.14)*	**14**	0.09	*(0.20)*	**14**	0.03	*(0.04)*	**15**	0.57^{NS}
severe schizophrenia	0.04	*(0.04)*	**16**	0.01	*(0.01)*	**16**	0.02	*(0.02)*	**16**	2.52^{NS}

[a] Analysis of variance with df=2.31; ** $p < 0.01$; * $p < 0.05$, NS $p \geq 0.05$.

panels on the weights assigned to the indicator conditions and (b) as the extent to which the rankings assigned to the weights by the three panels correspond.

The average weights of the three panels would not appear to be strongly diverging (see *table 3.3*). Univariate testing revealed that four weights differed significantly between the panels, on applying the Bonferroni inequality this number was reduced to three ($\alpha_F = \alpha/N$, in which $\alpha = 0.05$ and N = 16 (number of separate tests), therefore $\alpha_F = 0.05/16 = 0.003$). In all these cases the weight in the first panel differs from those in the other two panels. These differences can be explained by the course of the valuation procedure in the panels. For example, in the first panel, one of the panel members with a background in neonatology described the possible consequences of asphyxia at birth in explicit detail, invoking the image of a severely spastic child. This image did not arise during the other panel sessions. The difference between the panels was also investigated using multivariate analysis of variance; this was shown to be hardly to not at all statistically significant. (Wilks lambda = 0.13, p = 0.07; Hotellings T^2 = 5.08, p = 0.01).

Nor did the ranking of the weights assigned by the three panels vary much. The Spearman rank correlations between the three panels are high (*see table 3.4*), the ranking of the weights in the three panels is therefore very similar.

Based on these results, it may be concluded that the scale values derived at the group level are sufficiently reliable. These values were subsequently used for the disability scale in the interpolation procedure.

Table 3.4 – Inter-panel rank correlations (Spearman)

	panel 1	panel 2
panel 2	0.94	--
panel 3	0.95	0.97

3.2.3 Disability scale

The weights for the indicator conditions based on PTO yielded a disability scale on which the extremes are marked as the ‘best imaginable health state’ and the ‘worst imaginable health state’. The 16 values have been shown on the scale as the numerical value of the weights, with the corresponding standard deviation (*see figure 3.2*).

Figure 3.2 - Disability scale

3.3 Results of the interpolation session

All 38 panel members individually interpolated some 30 disease stages on the disability scale. The average time spent on this interpolation procedure was 1 hour and 15 minutes. The panel members indicated that the procedure was comprehensible and could easily be performed. When the results were inspected, all the interpolation forms proved to have been completed as instructed. The results of the interpolation session could therefore be calculated over all the panel members.

The panel members were explicitly invited to comment on the interpolation procedure. Most had no comments. The problem mentioned most frequently was the fact that the prognosis was not permitted to be factored into the weighting process. Only a single panel member reported difficulty with the difference in the order of the indicator conditions on the disability scale, based on the average weights from the panel sessions and his/her individual ranking of the indicator conditions.

3.3.1 Interpolation

The results of the weights for the disease stages in the interpolation procedure are summarized in *table 3.5*. The weights have been grouped into 11 so-called weight classes based on the scale values and a balanced distribution of the disease stages across the weight classes.

In summary, it may be concluded that:

- all acute infectious diseases are divided into the top two classes
- the neurological disorders occur only from class four on
- the lower two weight classes primarily contain the neurological and psychiatric disorders and the various cancers.

3.3.2 Reliability of the interpolations

The inter-rater reliability and the test-retest reliability of the weights for the disease stages assigned in the interpolation procedure could not be investigated for all disease stages, as not all of these had been interpolated by all panel members. Six disease stages constituted the common core which had been submitted to all 38 panel members in the interpolation procedure (*see table 3.6*). A similar reliability has been assumed for the reliability of the other disease stages in the interpolation procedure.

The degree of accordance between the individual panel members and the rest of the group was calculated to investigate the inter-rater reliability. The frequency of the Pearsonian correlations between weights of the individual panel members and the rest of the group (n=37) revealed a high degree of consensus between the members of the panels (*see figure 3.3*). This correlation was lower than 0.90 for three panel

Table 3.5 Disability weights for all disease stages, classified in 11 classes based on

the weights for the indicator conditions

Class	Disability weights	Code	Disease stage
1	1.00-0.99	34.2	periodontal disease (gingivitis)
		29.1P	acute nasopharyngitis (episode of 1 week in an otherwise healthy year)
		34.1	dental caries
		29.2P	acute sinusitis (episode of 2 weeks in an otherwise healthy year)
		1.1P	digestive tract infection, uncomplicated course (episode of 2 weeks in an otherwise healthy year)
		56.1	none to mild ADL limitations in elderly
		34.3	periodontal disease (pockets > 6 mm. deep)
		35.2P	acute urethritis (non STD) (episode of 1 week in an otherwise healthy year)
		4.1P	symptomatic acute gonorrhoea or Chlamydia trachomatis infection (episode of 1 week in an otherwise healthy year)
		35.1P	acute pyelitis/pyelonefritis (episode of 2 weeks in an otherise healthy year)
		31.1P	influenza (episode of 2 weeks in an otherwise healthy year)
		35.3P	acute cystitis (episode of 1 week in an otherwise healthy year)
		30.2P	acute bronchitis (episode of 2 weeks in an otherwise healthy year)
2	0.99-0.95	22.1	mild vision disorder (i.e., some difficulty reading small news-paper print, no difficulty recognizing faces at 4m. distance)
		29.2P	acute sinusitis (episode of 2 weeks in an otherwise healthy year)
		32.1P	active gastric or duodenal peptic ulcer (episode of 1 month in an otherwise healthy year)
		42.4	child in permanent stage after intentionally curative operation for pulmonary stenosis
		19.2	mild behavioural disorder (hyperactivity)
		42.1	young adult in permanent stage after intentionally curative operation for congenital atrial or ventricular septal defect
		1.2P	digestive tract infection, complicated course (episode of 2 - 4 weeks in an otherwise healthy year)
		28.1	mild to moderate asthma (symptom-free with or without main-tenance therapy)
		47.8	permanent impairment after luxation or distorsion of ankle or foot
		30.3P	acute bronchitis (more than one episode of 2 weeks per year)
		24.1	mild hearing disorder in elderly (i.e., some difficulty under-standing or actively participating in a conversation with one or more persons)
		34.4	edentulism
		50.1	basal cell skin cancer

Table 3.5 – continued

Class	Disability weights	Code	Disease stage
3	0.95-0.90	53.1	mild heart failure (NYHA 1-2)
		54.1	low back pain
		4.4	chronic hepatitis B infection without active viral replication
		47.6	permanent impairments after fracture of arm or shoulder
		36.2P	constitutional eczema (2 episodes of 6 weeks each of active eczema in an otherwise healthy year)
		13.1	uncomplicated diabetes mellitus
		50.2	squamous cell skin cancer, undisseminated
		26.1	mild stable angina pectoris (NYHA 1-2)
		17.5	mental retardation (IQ=70-84)
		30.1P	pneumonia (episode of 2 weeks in an otherwise healthy year)
4	0.90-0.85	52.1	epilepsy
		18.1	problem drinking (i.e., some physical, psychological or social problems caused by excessive alcohol intake)
		56.2	moderate to severe ADL limitations in elderly
		42.3	young adult in permanent stage after intentionally curative operation for Fallot's tetralogy or transposition of the great arteries
		23.1	mild to moderate congenital or early required hearing disorder
		51.3	mild to moderate agoraphobia
		4.2	late complications after gonorrhoeal or Chlamydia trachomatis infections (PID, subfertility)
		51.5	mild to moderate singular phobia
		24.2	moderate hearing disorder in elderly (i.e., some difficulty to understand or participate in a conversation with one person but great difficulties with conversations with more than one person)
		47.7	permanent impairmenties after fracture of leg or hip
		51.11	mild to moderate post traumatic stress disorder
		47.9	permanent impairmenties after burns
		16.1	mild depression
		39.1	osteoarthritis (grade 2) of hip or knee
		19.3	moderate to severe behavioural disorder (hyperactivity)
5	0.85-0.80	42.5	young adult in permanent stage after intentionally curative operation for pulmonary stenosis
		49.2	'remnant tuberculosis'
		51.1	mild to moderate panic disorder
		41.3	young adults with a low spina bifida aperta (sacral)
		51.7	mild to moderate social phobia
		2.2	permanent locomotor impairment after bacterial meningitis

Table 3.5 – continued

Class	Disability weights	Code	Disease stage
5 (cont'd)		51.13	mild to moderate diffuse anxiety disorder
		28.3	mild to moderate chronic obstructive pulmonary disease
		22.2	moderate vision disorder (i.e., great difficulty reading small newspaper print, some difficulty recognizing faces at 4m. distance)
		33.2	inflammatory bowel disease, in remission
		11.3	prostate cancer, clinically disease-free after primary therapy
		30.4	children with permanent impairment after moderate to severe bronchiolitis
		55.1	hip fracture, rehabilitation phase
		13.2	diabetes mellitus with neuropathy
		12.1	Non Hodgkin lymphoma of low-grade malignancy, dissemination stage I or II
		50.4	malignant melanoma I, no evidence of dissemination
		11.1	prostate cancer, accidentally detected localised prostate cancer, follow-up without active intervention ('watchful waiting')
		42.2	child/adolescent in permanent stage after intentionally curative operation for Fallot's tetralogy or transposition of the great arteries
		5.1	HIV seropositive
		8.2	colorectal cancer, clinically disease-free after intentionally curative primary therapy
6	0.80-0.70	15.1	schizophrenia (one psychotic episode, no permanent impairments)
		4.3	symptomatic non-fulminant acute hepatitis B infection
		38.1	mild rheumatoid arthritis
		23.2	severe congenital or early acquired hearing disorder
		51.9	mild to moderate obsessive/compulsive disorder
		2.3	permanent cognitive impairment after bacterial meningitis
		10.4	breast cancer, clinically disease-free after the first year
		10.1	breast cancer, diagnostic phase and primary therapy for non-invasive breast cancer or tumour < 2 cm
		14.1	mild dementia (only significant impairment of daily activities)
		11.2	prostate cancer, diagnostic phase and primary therapy for localised prostate cancer
		19.4	eating disorders (anorexia nervosa or bulimia nervosa)
		13.3	diabetes mellitus with nephropathy
		49.1	lung tuberculosis
		17.1	mild mental handicap (IQ=50-69)
		49.3	extrapulmonary tuberculosis

Table 3.5 – continued

Class	Disability weights	Code	Disease stage
7	0.70-0.60	5.2	AIDS-related complex
		4.6	compensated liver cirrhosis
		21.1	multiple sclerosis in ‘relapsing-remitting’ phase
		16.2	moderate depression
		45.1	children with permanent impairments after dysmature birth (‘small for gestational age’, birth weight < 5th percentile)
		53.2	moderate heart failure (NYHA 3)
		43.3	patient (10 - 40 jaar) with Down’s syndrome
		28.2	severe asthma (not symptom-free despite maintenance medication)
		45.3	children with permanent impairments after perinatal bacterial infection
		27.1	stroke, mild permanent impairments
		4.5	chronic hepatitis B infection with active viral replication
		6.2	oesophageal cancer, clinically disease-free after intentionally curative primarry therapy
		38.2	moderate rheumatoid arthritis
		24.3	severe hearing disorder in elderly (i.e., great difficulty or unable to understand or participate in a conversation with one other person)
		47.1	permanent impairments after mild skull/brain injury
		7.2	stomach cancer, clinically disease-free after intentionally curative primarry therapy
		33.1	inflammatory bowel disease, active exacerbation
		50.3	squamous cell skin carcinoma with lymph node dissemination
8	0.60-0.50	51.6	severe singular phobia
		39.2	osteoarthritis (grade 3 - 4) of hip or knee
		22.3	severe vision disorder (i.e. unable to read small newspaper print, great difficulty or unable to recognize faces at 4m. distance)
		8.1	colorectal carcinoma, diagnostic phase and primary therapy
		50.5	malignant melanoma II, lymph node dissemination, no distant dissemination
		17.2	moderate mental handicap (IQ=35-49)
		9.1	lung cancer, diagnostic phase and primary therapy for operable non small-cell lung cancer
		45.4	children with permanent impairments after perinatal viral infection
		9.3	lung cancer, clinically disease-free after primary therapy for non small-cell lung cancer

Table 3.5 – continued

Class	Disability weights	Code	Disease stage
8 (cont'd)		20.1	initial stage M. Parkinson (initially unilateral, later bilateral tremors and rigidity; slowness, impaired swallowing and speech; disturbance of equilibrium; patients are able to function indepedently)
		44.1	children with permanent impairments 5 years after premature birth (< 32 weeks),
		45.2	children with permanent impairments after asphyxia (APGAR < 7 after 5 minutes)
		41.2	young adults with medium level spina bifida aperta (L3 to L5)
9	0.50-0.35	43.2	child, age below 10. with Down's syndrome, without other congenital anomalies
		51.12	severe posttraumatic stress disorder
		28.4	severe chronic obstructive pulmonary disease
		7.1	cancer of the stomach, diagnostic phase and primary therapy
		9.7	small-cell lung cancer, clinically in remission
		18.2	manifest alcoholism (severe social problems caused by excessive alcohol intake)
		19.1	autism (i.e., qualitative deficits in social interactions and communication)
		12.3	Non Hodgkin lymphoma of intermediate/high malignancy grade, dissemination stage I
		51.4	severe agoraphopia
		6.1	oesophageal cancer, diagnostic phase and primary therapy
		51.10	severe obsessive/compulsive disorder
		5.3	AIDS, first stage
		26.2	severe stable angina pectoris (NYHA 3-4),
		47.4	paraplegia
		51.8	severe social phobia
		51.14	severe diffuse anxiety disorder
		12.2	Non Hodgkin lymphoma of low malignancy grade, dissemination grade III-IV
		27.2	stroke, moderate permanent impairments
		14.2	moderate dementia (independent living living is not possible without limited supervision)
		11.4	prostate cancer, disseminated
		43.4	adult, over 40 years of age, with Down's syndrome
		53.3	severe heart failure (NYHA 4)
		56.3	elderly with extreme ADL limitations or complete ADL dependence

Table 3.5 – continued

Class	Disability weights	Code	Disease stage
10	0.35-0.20	21.2	multiple sclerosis in primary or secundary progressive phase
		9.6	small-cell lung cancer, diagnostic phase and chemotherapy
		41.1	young adult with high level spina bifida aperta (L2 or higher)
		10.2	breast cancer, diagnostic phase and primary therapy for breast tumour 2-5 cm. and/or local lymph node dissemination
		51.2	severe panic disorder
		43.1	child, age below 10 with Down's syndrome, with other congenital anomalies
		15.2	schizophrenia, several psychotic episodes, some permanent impairments
		42.6	child/adolescent in permanent stage with complex not curatively operable congenital heart disease
		47.2	permanent impairments after moderately severe skull/brain injury
		7.3	cancer of the stomach, irradically removed or disseminated
		47.3	permanent impairments after severe skull/brain injury
		12.4	Non Hodgkin lymphoma of intermediary/high grade malignancy, dissemination stage II, III of IV
		2.4	permanent cognitive and locomotor impairment after bacterial meningitis
		17.4	extreme mental handicap (IQ<20)
		16.3	severe depression, without psychosis
		9.2	non small-cell lung cancer, diagnostic phase and primary therapy for inoperable cancer
		10.5	breast cancer, disseminated
		20.2	intermediate stage M. Parkinson (swallowing and speech severely impaired; autonomic nervous system disturbances; patients are ADL-dependent, but are able to move without help)
11	0.20-0.00	50.6	malignant melanoma III, disseminated
		15.3	schizophrenia, several psychotic episodes, obvious permanent impairments
		10.3	breast cancer, diagnostic phase and primary therapy for locally advanced breast cancer (tumour > 5 cm)
		17.3	severe mental handicap (IQ=20-34)
		8.3	colorectal cancer, irradically removed or disseminated
		16.4	severe depression with psychosis
		18.3	psycho-organic disorder (delirium) caused by excessive alcohol intake
		47.5	tetraplegia
		4.7	decompensated liver cirrhosis
		6.3	oesophageal cancer, irradically removed or disseminated

Table 3.5 – continued

Class	Disability weights	Code	Disease stage
11 (cont'd)		9.4	non small-cell lung cancer, disseminated
		27.3	stroke, severe permanent impairments
		20.3	end-stage M. Parkinson (wheelchair and bed patient, severely handicapped)
		0.5	end stage disease otherwise unspecified
		38.3	severe rheumatoid arthritis
		14.3	severe dementia (permanent supervision required)
		15.4	schizophrenia, several psychotic episodes, severe and increasing permanent impairments

members only, while the lowest correlation was 0.73. The average correlation of the weights each time between an individual panel member and the rest of the panel (comparable to item-rest correlation) was 0.95.

The test-retest reliability for the six common core disease stages was calculated for 33 of the 38 panel members. Three panel members stated that they did not wish to be asked to participate in a test-retest study, two members of the panels failed to return the retest forms after being sent a reminder. The results showed that the average disability weights for the six common core disease stages for the interpolation session and the retest after two months barely differed, while the ranking correlations per disease stage between interpolation and retest were moderate to low (*see table 3.6*). The correlation found between the interpolations of all common core conditions from the interpolation procedure and the retest was 0.94. There were, apparently, intra-individual movements in the interpolation of disease stages, while the *average* disability weights were stable. There may be a context effect: during the interpolation procedure, the common core conditions were presented between the other interpolation conditions. The intra-individual movements did not occur systematically in a certain direction, however, and the stability of the average weights was excellent, so that the results are usable.

3.4 Validity of the weights

The validity of the weights was studied firstly by comparing them to the disability weights elicited in Murray's GBD study. Secondly, the disability weights were mutually compared within and between diseases. Thirdly, a comparison was made with theoretical weights derived from a statistical model.

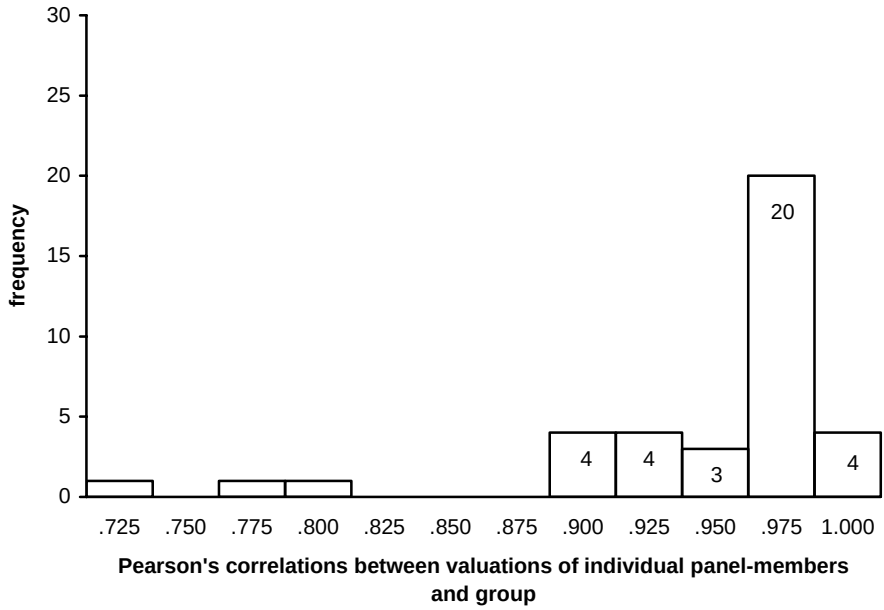


Figure 3.3 - Reliability (individual-rest correlations) of interpolations of the common core conditions

Table 3.6 – Test-retest reliability of disability weights for 'common core' disease stages

disease stage	interpolation		retest		
	mean	median	mean	median	r_S [a]
severe hearing disorders in elderly	0.63	0.66	0.64	0.66	0.23
mild to moderate post traumatic stress disorder	0.87	0.86	0.85	0.86	0.28
severe heart failure	0.35	0.40	0.32	0.27	0.65
end-stage disease otherwise unspecified	0.07	0.04	0.08	0.04	0.62
multiple sclerosis ('relapsing-remitting')	0.67	0.66	0.62	0.66	0.62
influenza (duration 2 weeks, in annual profile)	0.97	0.99	0.99	0.99	0.34

[a] r_S = Spearman rank correlation

3.4.1 Comparison with the weights from the GBD study

The weights derived from the disease stages in the present study were compared to the disability weights assigned the indicator conditions in the GBD study, to the extent that similar health states were concerned (*table 3.7*). Murray provides 'severity weights' for these 'indicator conditions', divided into 7 'disability classes'. The Dutch disability weights were therefore also divided into similar classes for the sake of comparison. Of the 22 indicator conditions in the GBD study, 12 had a comparable counterpart in the Dutch study (Murray's other 10 indicator conditions were not included in the Dutch disability weights study because they did not comply with the criteria for inclusion in the Public Health Status and Forecast 1997 list of diseases).

The results showed that the weights derived in both studies corresponded rather well. Five disease stages proved to have been classified into the same class in both studies, two other disease stages were situated virtually on the border between two disability classes. The other five disease stages ended up either one class higher or lower. These differences are partially explainable by the difference in the context of the valuations (global versus the Netherlands). Infertility and mental retardation probably have less far-reaching consequences in the Dutch situation than in developing countries. Angina pectoris and depression were nonetheless the sole disease stage submitted as indicator conditions for weighting for these diagnoses in the GBD study. In the Dutch weights study, various disease stages were included for these diagnostic groups and the complete disease model was shown. Hence, severe depression may have been more heavily weighted because of the fact that moderate and mild depression were also included. Analogous to this is the fact that mild stable angina pectoris may also have possibly been weighted more lightly due to the inclusion of severe stable angina pectoris.

All in all, these results support the validity of the weights.

3.4.2 Comparison of disability weights per disease

In order to judge the extent to which the weights derived are plausible, first the weights were systematically compared *per disease*. In most cases, it was possible to rank per diagnostic group the stages according to severity. An infectious disease of the digestive tract running an uncomplicated course is obviously less severe than one running a complicated course. If this same order is reflected in the weights, this offers an indication in favour of the validity of these weights.

A systematic comparison of the weights per disease revealed that the order of the weights corresponded in virtually all cases with the logical order. Hence at the level of the ranking, it was concluded that the weights were valid. There were three exceptions, where an unexpected rating had appeared, one of which will be discussed as an illustration. The stage 'compensated liver cirrhosis' in hepatitis B had been assigned a weight indicating that this was less severe than the stage 'chronic carrier with active

viral replication'. This (statistically insignificant) reversal of the order may be explained by the label ('compensated' perhaps sounds less threatening than 'chronic with active replication'), possibly combined with a relative unfamiliarity of the participants with the symptoms occurring in chronic liver diseases.

3.4.3 Comparisons of disability weights between diseases

If a logical ranking can be assigned according to a severity scale *between diseases*, theoretically this should offer a second means to judge the validity of the derived weights. The possibilities of any such a priori ordering of conditions with dissimilar sequelae are limited: it was, in fact, this task of assigning such an order with which the participants were charged. Nonetheless, various possibilities do arise, such as a comparison of the weights between more or less similar diseases, e.g. between various types of

Table 3.7 – Comparison GBD - Dutch disability weights

GBD (WHO/Worldbank)			Dutch study		
'indicator condition'	'disability class'	'severity weight'	disease stage	disability class	disability weight[a]
infertility	3	0.12 - 0.24	late complications after STD infection	2	0.11
angina	3	0.12 - 0.24	mild stable angina pectoris	2	0.08
rheumatoid arthtris	3	0.12 - 0.24	mild rheumatoid arthritis	3	0.21
deafness	4	0.24 - 0.36	severe hearing disorder in elderly	5	0.37
blindness	6	0.50 - 0.70	severe vision disorder	5	0.43
mild mental retardation	5	0.36 - 0.50	mild mental handicap	4	0.29
Down's syndrome	5	0.36 - 0.50	Down's syndrome without comorbid conditions	6	0.51
paraplegia	6	0.50 - 0.70	paraplegia	6	0.57
unipolar major depression	6	0.50 - 0.70	severe depression	7	0.76
active psychosis	7	0.70 - 1.00	severe schizophrenia	7	0.98
dementia	7	0.70 - 1.00	severe dementia	7	0.94
quadriplegia	7	0.70 - 1.00	tetraplegia	7	0.86

[a] Dutch disability weights in the direction analogous to Murray's disability weights (1-disability weight)

cancer. To illustrate, a number of the assumed rankings according to severity of the disease are summed up below:

- diseases of the upper respiratory tract, increasing in severity: common cold (1 episode of 1 week in an otherwise healthy year), acute bronchitis (1 two-week episode in a year), acute bronchitis (more than 1 two-week episodes a year).
- infectious diseases: HIV/AIDS worse than all other
- cancers of the digestive tract: stage of diagnosis and primary therapy for carcinoma of the oesophagus and stomach of similar severity (i.e. we expected approximately the same disability weight); both more severe than the same stage of colorectal cancer (inter alia because of a better prognosis of colorectal cancer).
- all types of cancer: approximately the same weights were expected for all disseminated stages
- heart disease: for severe stable angina pectoris (NYHA 3) a weight was anticipated in the same order of magnitude as for severe stable heart failure (NYHA 3).
- cognitive disorders: similar weights were expected for 'moderate mental retardation', 'child with Down's syndrome without co-morbidity' and 'adult with Down's syndrome'.

On comparing the assumed rankings with the final weights, few discrepancies were revealed. Where discrepancies turned up, these could be explained, e.g. by the way conditions had been presented to the panel. In this way 'severe stable angina pectoris' had been assigned a more severe weight than 'severe stable heart failure'. Closer consideration revealed that severe stable angina pectoris had been presented as 'NYHA 3-4' and severe stable heart failure as 'NYHA 3', which readily explained the difference in assigned weight. To recapitulate, this systematic mutual comparison of weights pointed in favour of the validity of the weights.

Reviewing the list of disability weights as a whole brings a few other striking points to light. In the first place, the diseases presented within the framework of an annual profile ('a one- or two-week episode in an otherwise healthy year) on the whole were not rated as severe, even where conditions were concerned which were less easily dismissed, such as 'infectious diseases of the digestive tract with a complicated course', 'acute pyelonephritis', and 'more than one episode of acute bronchitis'. Although perfect recovery is probable in these cases, some differentiation in the weights assigned compared to e.g. the common cold was expected. It should be noted that a calculable minimum applies to the weights in these annual profiles, assuming perfect linearity. For an episode of # weeks spent in the worst imaginable health state (weight 0), followed by a perfect recovery, this is:

$$\text{weight annual profile} = \frac{(\#\ \text{weeks}*0)+(52-\#\ \text{weeks}*1)}{52}$$

For an episode with a duration of 1 week in an otherwise healthy year, this works out to 0.98; for a two-week episode, 0.96; for a four-week episode, 0.92. A lower weight would mean that the health state in the shorter episode would be assigned a weight smaller than 0, or that the given duration of the episode was unrealistic. Pneumonia proved to be a case in this point. This example serves to illustrate how important a precise estimation of the duration of a short episode of disease is.

In the second place, some results would appear to have been the victim of an attitude of 'unknown is unloved', or in other words, of a tendency of the participants to rate a disease about which they know relatively little as relatively severe. This was possibly the case for e.g. Non Hodgkin lymphoma, hepatitis B, chronic inflammatory bowel disease and TBC.

3.4.4 Assessing validity of the weights using EuroQol 5D+ classifications

As described in *section 2.2.3* a description of each disease stage to be valued was added in EuroQol 5D+ terms. The extended EuroQol (5D+ variant) comprises six dimensions, each with three levels. Using this system, the health states could be coded according to the classification shown in *table 3.8*. For example: a functional health state for 'severe vision disorder',

Table 3.8 – The EuroQol 5D+ classification for health status

dimension	level	code
mobility	no problems in walking about	1
	some problems in walking about	2
	confined to bed	3
self-care	no problems with washing or dressing self	1
	some problems with washing and dressing self	2
	unable to wash or dress self	3
USUAL activities	no problems with performing usual activities *(e.g. work, study, housework, family or leisure activities)*	1
	some problems with performing usual activities	2
	unable to perform daily activities	3
pain/discomfort	no pain or discomfort	1
	moderate pain or discomfort	2
	extreme pain or discomfort	3
anxiety/depression	not anxious or depressed	1
	moderately anxious or depressed	2
	extremely anxious or depressed	3
cognition	no problems in cognitive functioning *(e.g., memory, concentration, coherence, IQ)*	1
	some problems in cognitive functioning	2
	extreme problems in cognitive functioning	3

characterized by:
- no problems in walking about
- some problems with washing and dressing self
- unable to perform usual activities
- no pain or discomfort
- moderately anxious or depressed
- no problems in cognitive functioning

can be coded as 123121: code 1 (no difficulty) on the first dimension (mobility), code 2 (some difficulty) on the second dimension (self care), etc.

On the basis of this EuroQol 5D+ coding, a statistical model could be used to estimate a theoretically expected valuation for each disease stage. This would not necessarily have to correspond in either value or ranking with the empirically determined valuation. In the present study, complete agreement between the *estimated* and the *derived* weights was not expected, as the (in particular, prognostic) information of the diagnostic label has not or not fully been included in the EuroQol 5D+ descriptions. Nonetheless, comparison between the weights derived and those estimated by the statistical model can be meaningful. Differences in ranking should be able to be interpreted.

Valuations can be estimated for the standard EuroQol version with five dimensions with the help of the Busschbach model, a regression model based on empirical valuations of a limited number of health state descriptions. (Busschbach, 1997) This model was not used by us for this purpose because:
- in the Dutch disability weights study, use was made of a more elaborate EuroQol 5D+ version to which a sixth dimension was added. It was not possible to calculate a regression weight for this sixth dimension. Deriving this from the existing five-dimensional model would entail too many uncertainties.
- the original non-standardized Busschbach model yielded a spread of estimated values of between 3 and 77 only (instead of the theoretical spread of 0 - 100), and standardization according to the model led to insufficiently reliable regression weights. The disability weights elicited have a spread reaching from 4 to 100, which means that a comparison must be made with estimated weights able to cover the entire reach.

Instead, a simple additive model was used, assuming equal weights for all dimensions and equidistance between the levels:

$$\mathrm{V}(Q) = 100 - (\sum_{i=1}^{6} x_i - 6) \times \frac{100}{12}$$

in which:
V(Q) = valuation V for EuroQol 5D+ health state Q;
x_i = category x of dimension i, $x \in \{1.2.3\}$.

In this model, state 111111 is assigned a valuation of 100; the state 112111 a valuation of 91.67; state 1211212 a valuation of 73; state 222222 a valuation

of 50; health state 333222 a valuation of 25 and 333333 a valuation of 0. The vision disorder example, 123121 is assessed at 67 according to this model.

A number of disease stages were characterized by two or more functional health states and hence by different EuroQol 5D+ descriptions, with a prevalence distribution in percentages. When calculating V(Q), the valuations were calculated in the same way as the weighted assessment of the various functional health states.

$$\mathrm{V}(Q) = \left[a\% * (100 - (\sum_{i=1}^{6} x_i - 6) \times \frac{100}{12}) \right] + \left[b\% * (100 - (\sum_{i=1}^{6} x_i - 6) \times \frac{100}{12}) \right]$$

Disease stages have also been included which have been valued only as a stable residual state with disabilities; in such cases, the healthy state was not included in the weighting process.

The average difference between the derived and the estimated weights (excepting the annual profiles) was -0.10 (SD = 0.13; median = -0.08), i.e. the empirically derived weights were on average lower than the weights derived from the statistical model. On comparing the ranking of the derived and the estimated weights, a good correlation was found (Spearman rank correlation rs = 0.81). The differences in ranked order may be explained by the addition of a label of a disease for which the prognosis is poor, such as AIDS or cancer, which are rated much worse than is assumed on the basis of the model. Moreover, the comparison between the estimated and the derived weights differs greatly where disease stages are concerned with a poor cognitive dimension of the functional health state, e.g. ‘dementia’, ‘mental handicap’. This points to the failure of the additive model to represent properly the importance of the cognitive dimension. This finding corresponds with the results of the comparison between the EuroQol 5D and 5D+ versions. (Krabbe, 1997b)

3.5 Results lay panel

The lay panel was composed of 7 members, 4 men and 3 women, who took part in a panel session. The average age was 39 (minimum 24, maximum 64). None of the panel members worked in medicine and their health care experience was limited to their own illnesses and/or illnesses suffered by family members.

All the participants in the lay panel carried out the valuation exercises for the 16 indicator conditions. They tended to ask for more details about the clinical picture corresponding to an indicator condition than the physicians had done. It was striking to note that they sought to estimate the relative severity of the disease stages to a far greater degree based on the standardized descriptions of the health states associated with the indicator conditions. The panels of medical experts used this information primarily for a more precise specification of the disease stage described while the lay

panel needed this information simply in order to form a picture of a patient with the disease stage described.

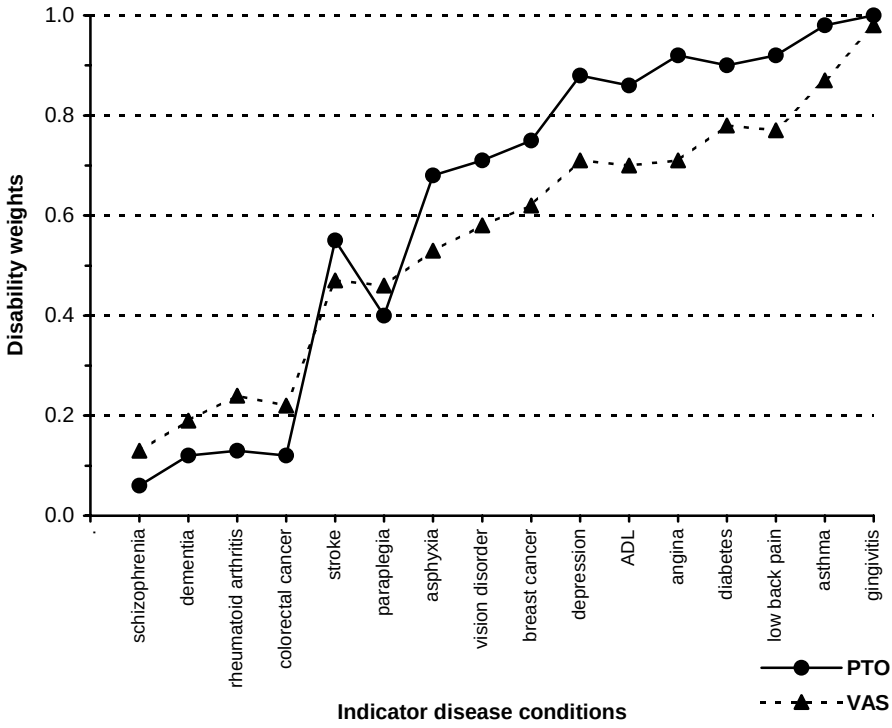


Figure 3.4 - Weights lay panel based on average valuations

When asked, they indicated that dimensions of ‘pain/discomfort’, ‘anxiety/depression’ and ‘cognition’ primarily played a role in coming to a valuation. On the other hand, that (perceived) prognosis could play a role in the ultimate valuation of a health state was never an issue for the lay panel. Among the medical experts, the fact that only any uncertainty regarding the prognosis, but not the prognosis itself was allowed to be part of the weighting process was, on the contrary, a recurrent point of discussion.

In the end, the person trade-off method was accepted by all participants as a valuation method. The majority reported finding PTO2 conceptually more difficult than PTO1, and in actual fact may have only applied PTO1. Only a single participant admitted to having used PTO2 in valuing ‘severe diseases’.

The results of the lay panel session are shown in figure 3.4. The ranking assigned by the PTO and VAS valuations were virtually alike. The average PTO valuations for CVA and paraplegia by the lay panel were opposite to the VAS valuations. The Spearman rank correlation between PTO and VAS was 0.98. The absolute PTO valuations at either end of the figure were more extreme than the VAS valuations. This agrees with the results of the panels of medical doctors. In *table 3.9*, the absolute values of PTO and VAS valuations do vary somewhat, but less than for the panels of medical experts. However, the number of participants in the lay panel was not large which may bias the results of the test. During nonparametric testing (Wilcoxon) more significant differences were revealed between PTO and VAS.

Comparing the weights based on average PTO valuations assigned by the lay panel to those assigned by the panels of medical experts, few differences were seen (*see figure 3.5*). The valuations diverged for the indicator conditions 'stroke', 'asphyxia' and 'vision disorder', although these differences were not statistically significant.

Based on the above results, the conclusion that it makes little difference

Table 3.9 – Disability weights lay panel: PTO and VAS (mean, standard deviation, median) (N=7)

disease stage	PTO1/PTO2			VAS			
	mean	*SD*	median	mean	*SD*	median	Student *t* [a]
periodontal disease (gingivitis)	1.00	*0.00*	1.00	0.98	*0.03*	0.99	1.75^{NS}
mild/moderate asthma	0.98	*0.01*	0.98	0.87	*0.14*	0.94	2.20^{NS}
low back pain	0.92	*0.06*	0.95	0.77	*0.15*	0.75	2.91*
diabetes mellitis (uncomplicated)	0.90	*0.18*	0.95	0.78	*0.27*	0.90	2.63*
mild angina pectoris	0.92	*0.06*	0.95	0.71	*0.18*	0.70	3.11*
moderate ADL-limitations	0.86	*0.15*	0.91	0.70	*0.15*	0.70	3.77**
mild depression	0.88	*0.06*	0.91	0.71	*0.08*	0.70	3.21*
breast cancer (clinically disease free)	0.75	*0.19*	0.80	0.62	*0.11*	0.62	2.68*
severe vision disorder	0.71	*0.17*	0.77	0.58	*0.22*	0.60	2.05^{NS}
state after asphyxia	0.68	*0.16*	0.71	0.53	*0.12*	0.50	2.55*
paraplegia	0.40	*0.19*	0.44	0.46	*0.09*	0.49	-0.66^{NS}
state after stroke, moderate permanent impairments	0.55	*0.23*	0.50	0.47	*0.16*	0.50	2.03^{NS}
colorectal cancer (disseminated)	0.12	*0.09*	0.13	0.22	*0.19*	0.18	-2.07^{NS}
severe rheumatoid arthritis	0.13	*0.06*	0.13	0.24	*0.17*	0.20	-1.63^{NS}
severe dementia	0.12	*0.09*	0.15	0.19	*0.09*	0.20	-1.79^{NS}
severe schizophrenia	0.06	*0.07*	0.04	0.13	*0.12*	0.10	-2.42*

[a] Paired t-test with df=6; ** = $p < 0.01$; * = $p < 0.05$; NS = $p \geq 0.05$.

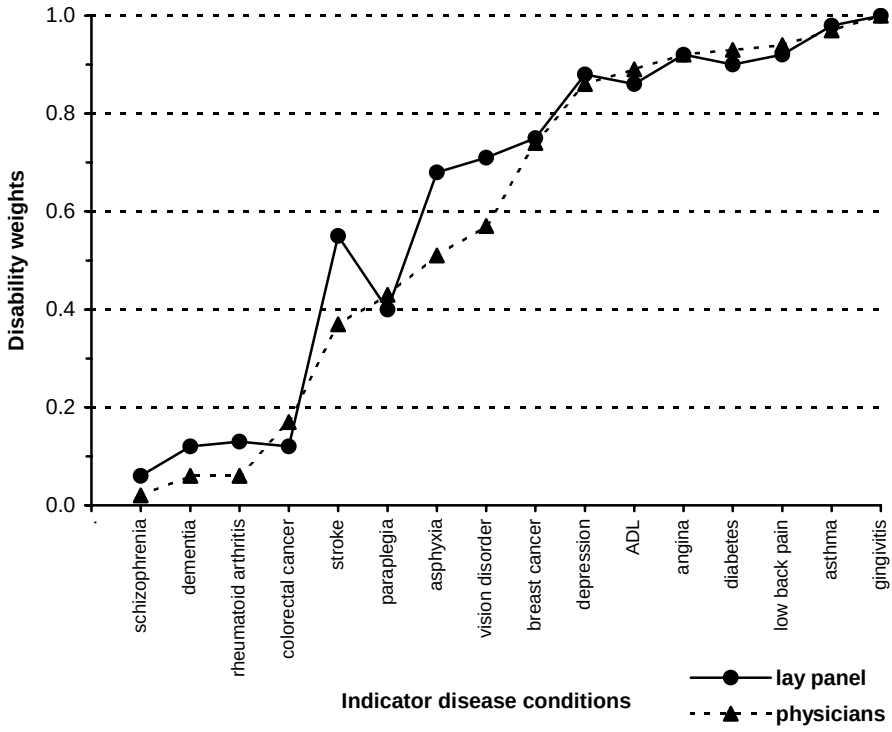


Figure 3.5 - Comparison PTO valuations panels of medical doctors and lay panel

whether the panel is composed of health care workers or people with no medical practice or experience whatsoever is warranted. After all, in the end the valuations differ hardly at all. Froberg and Kane come to this same conclusion in their review. (Froberg, 1989) The way in which these valuations are arrived at does seem to be different. Without the addition of the functional health state description, it would hardly be possible for laymen to value the naturalistic descriptions. And the considerations on which the final assessments were based also differ between panel members with and without medical knowledge/ experience.

4

CONCLUSIONS AND RECOMMENDATIONS

THE OUTCOME of the Dutch project on 'Disability weights for diseases' is a coherent set of disease-specific disability weights for 175 disease stages, derived from the 52 diseases selected in the Public Health Status and Forecast 1997 study *(VTV-97)*. This demonstrates first of all that it is possible to derive reliable weights for a large number of different diseases in a reasonable period of time. A second important point was that the weights were determined in a comprehensive approach. In other words, these 175 disease stages were all weighted on the same scale. When applying the weights so obtained in the calculations on the burden of disease for the different diseases, the results will be more mutually comparable than was hitherto the case because of this mutual coherence between the weights (and naturally on the condition that the combination with mortality data occurred uniformly). This will enable insight to be gained into, for example, the share of specific diseases in the total burden of disease. This is important in order to be able to identify the key points on which to build the policy on public health.

This section will first examine the possible uses for the disability weights elicited in the present study. These will be followed by a number of research recommendations on topics which shape the current possibilities for application and could further expand these possibilities in the future.

4.1 Possible uses

The disability weights are tied to no specific method of combining morbidity and mortality data. With the help of the disability weights derived, it should in principle be possible to calculate weighted Healthy Life Expectancies, DALEs (disability adjusted life expectancies), DALYs (disability adjusted life years) and QALYs (quality adjusted life years).

The disability weights are suitable for a broad spectrum of applications in public health research and health services research. Examples of their application in public health research are estimations of the total burden of

disease in the population and of the share accounted for by certain disease groups. Such calculations have been performed within the scope of the Public Health Status and Forecast 1997 study. The results are presented in the report on these surveys (Ruwaard, 1997). They are important for defining the key points for public health policy. If the figures on disease prevalences are sufficiently valid, trends over time in the total burden of disease can also be described together with the increase or decrease, whichever the case, of the share of specific diseases in the total burden of disease. And finally, the weights are suitable for international comparisons of the burden of disease, on the condition that the international transferability of the weights has been sufficiently demonstrated and that the necessary epidemiological data are available.

In health services research, the weights may be valuable for use in efficiency initiatives, such as the assessment of pharmaceuticals (pharmaco-economic research) and other medical interventions. Composite health outcome measures such as the DALY and QALY are preeminently suitable for comparing the effects of dissimilar facilities for different types of disorders, such as was recently recommended by the Scientific Council for Government Policy in the Netherlands (WRR). The application of this set of weights in health technology assessment studies (MTA) on medical facilities can render the results of such studies more mutually comparable (at least in respect of the quantification and valuation of outcomes), so that for example QALY league tables may become more meaningful.

Use of the same set of disability weights in public health and health services research can foster the integration of the information obtained from the two fields. The findings from assessment studies on separate interventions can be related to effects on public health as a whole. These disability weights could consequently become an important element for generating the information needed on which to base public health policy decisions.

The assessment of the consequences of the introduction of thrombolytic agents may serve as an illustration of the potential added value of applying one and the same set of coherent disease-specific disability weights. The cost effectiveness ratio of thrombolytic drugs can be determined through an economic assessment in terms of guilders per DALY avoided or QALY gained. A routine application of these drugs in patients presenting with an acute myocardial infarction will cause a change in the burden of disease at the population level: death is deferred, but the incidence of heart failure and stroke will increase. This change can be captured with the help of the disability weights and epidemiological data in absolute figures on the DALYs avoided (or QALYs gained). Reasoning thus, using the set of disease-specific disability weights now available the cost effectiveness ratios for a variety of interventions can be related to the total costs and effects on the population’s health. Insights into such relations are vital for policy decisions relating to prioritization of services.

4.2 Research recommendations

The derivation of these disability weights is an important step towards a more integral description of health and disease. A responsible application in the various areas mentioned will require, however, a better epidemiological database, and more precisely defined descriptions of the disease stages to be valued. Moreover, a study should also be made of the international transferability of the disability weights derived. These research recommendations are further elaborated in the following. After all, disability weights are not an end in themselves, but simply a link in the complex set of data used to form a picture - in standardized terms - of the health state of a population.

4.2.1 The need for corresponding epidemiological data

The usability of the weights in public health research will depend primarily on the availability of consistent and comprehensive epidemiological data on the diseases in question. The stages for which the weights were derived must fit with the epidemiological data, e.g. a known average or median duration of each disease stage. The 'list of diseases and disease stages' is therefore an essential element of the project. On the one hand, the disease stages distinguished must be homogeneous as regards health status, treatment and prognosis, in order to present those making the assessment with a uniform state to be valued. On the other hand, epidemiological data must be available for precisely these disease stages. This is an important point which merits additional research.

4.2.2 Standardized description of each disease stage

As an extension of the list of diseases and disease stages, a standardized representative description of the functional health state in each stage was given. The addition of such a description couched in EuroQol 5D+ terms proved to be indispensable to the valuation process in the Dutch disability weights study. Whether or not the EuroQol 5D+ descriptions applied in the present study are indeed accurate and representative requires further investigation. Improving the accuracy may require refinement of the classification used to describe the functional health states, for example into 5 levels instead of 3 per dimension. Mutually comparable empirical data that document the health states associated with a large number of diseases in average (i.e., non-academic) treatment settings are helpful to enhance representativeness of the health state descriptions.

4.2.3 Reliability and validity of the current disability weights

The study described has hitherto yielded positive indications regarding the reliability and validity of the weights elicited. It should be added that in particular the validity of the valuations tends to be generally hard to establish. The Global Burden of Disease study is in fact the only study with

which the results of the Dutch project can be compared, up to a certain point. As described in *section 3.4.1*, the results of that comparison are encouraging. The disability weights obtained could only be compared within the study itself, as described in *sections 3.4.2 and 3.4.3*. Future research must be directed at the validity and reliability of the procedures followed. This could include:

- research into the properties of PTO (test-retest reliability; validity, through formal empirical comparison with other valuation methods; etc.)
- research into the effect of presenting a PTO with a life of 1 year followed by death versus a more realistic profile of a life conforming to the life expectancy of the group described.
- research into the effects of the choice of indicator conditions on the final weights.
- closer comparative research into the effects of the background of the individuals making the assessment (the medical versus lay panel)
- research into the presentation of disease stages [described in a naturalistic manner (diagnostic labels) or as generic health state descriptions; or by other, e.g. more visual presentations], in connection with the different kinds of panels.

4.2.4 Trends in disability weights

In due course, disease-specific disability weights require investigation of trends in the weights over time. Due to the development of e.g. new treatment methods, it is not improbable that the sequelae of certain conditions lessen in severity. An example is a HIV infection. Several months after the disability weights for HIV and AIDS had been empirically determined, it was announced that certain cocktails of new drugs could yield considerable improvement in the prognosis. AIDS could thus in time be reduced to a ‘true’ chronic disease instead of a dragged-out process leading in the medium term to certain death. The weights for AIDS and related disease stages derived before this new information was published may perhaps assign too high a weight to the sequelae following a HIV infection.

4.2.5 Disaggregation of disability weights for application in assessment studies of specific interventions

In the foregoing, it was argued that application of the disability weights derived in economic assessment studies can in principle lead to the results of such studies becoming mutually more easily comparable, while the results can moreover be related to public health as a whole. The disability weights derived in this study encompass a broad spectrum of diseases. The flip side of this comprehensiveness is a certain degree of crudeness, as a result of which the current disability weights may possibly not be sufficiently refined for use in the economic evaluation of specific medical technologies. This can be redressed by disaggregation, as it were, of the disability weights now available. For example, the weight for ‘epilepsy’ as determined in the present study, is actually composed of the combination of weights for different types

of epilepsy. However, the disability weights for the various different types of epilepsy can be individually determined for the purpose of use in economic analyses. If the same methods are followed as were used in the present study, the commensurability towards the higher aggregation level (mutual studies, relation to public health) will be maintained.

4.2.6 Co-morbidity

In the study contained in this report, hardly any account was taken of the consequences of the simultaneous occurrence of more than a single condition in an individual (co-morbidity). This amounts to the assumption that the disability weight for a combination of conditions equals the sum of the disability weights of each of the components of the combination. Reality is probably more complex. Empirical research is needed to assess the combined effects of more than one condition on disability.

4.2.7 International comparison and application

Although the worldwide scale of the Global Burden of Disease study is too expansive in some respects for the Dutch situation, the Netherlands is not an autonomous island. The applicability of the weights should be studied in a European context. A European research network in this area, made up of Great Britain, France, Spain, Sweden, Norway, Denmark and the Netherlands has been granted funds by the EU within the scope of the BIOMED-II program to launch an international study (starting in 1998) into the similarities and differences between the countries of Western Europe in respect of disability weights for disease and the possibilities which these offer for making calculations on the burden of disease.

REFERENCES

Brooks R. EuroQol: the current state of play. Health Policy 1996;37:53-72.

Busschbach JJV, McDonnell J, Essink-Bot ML, Hout BA van. Estimating a parametric relation between health description and health valuation using the EuroQol instrument. Submitted for publication, 1997.

Essink-Bot ML. Health status as a measure of outcome of disease and treatment. Thesis, Erasmus University Rotterdam. Rotterdam: 1995.

EuroQol Group. EuroQol - a new facility for measurement of health-related quality of life. Health Policy 1990;16:199-208.

Froberg DG, Kane RL. Methodology for measuring health-state preferences I-IV. Journal of Clinical Epidemiology, 1989;42:345-54; 459-71; 585-92; 675-85.

Krabbe PFM, Essink-Bot ML, Bonsel GJ. The comparability and reliability of five customary health-state valuation methods. Social Science and Medicine 1997;45(11): 1641-52(a)

Krabbe PFM, Stouthard MEA, Essink-Bot ML, Bonsel GJ. The effect of adding a cognitive dimension to the EuroQol multi-attribute health-status classification system. Submitted for publication. 1997(b)

Murray CJL, Lopez AD (ed). Global Burden of Disease and Injury series. Vol. 1: The Global Burden of Disease. Boston: Harvard University Press, 1996. (a)

Murray CJL. Rethinking DALYs. Chapter 1 (pp. 1-98) in: Murray CJL, Lopez AD (ed). Global Burden of Disease and Injury series. Vol 1: The Global Burden of Disease. Boston: Harvard University Press, 1996. (b)

Murray CJL, Lopez AD. Regional patterns of disability-free life expectancy and disability adjusted life expectancy: Global Burden of Disease study. Lancet 1997;349:1347-52. (a)

Murray CJL, Lopez AD. Global mortality, disability and the contribution of risk factors: Global Burden of Disease study. Lancet 1997;349:1436-42. (b)

Nord E. The person trade-off approach to valuing health care programs. Med Decis Making 1995;15:201-8.

Olsen JA. Person vs years: two ways of eliciting implicit weights. Health Economics 1994;3:39-46.

Patrick DL, Bush JW, Chen MM. Methods for measuring levels of well-being for a health status index. Health Services Research 1973;8:228-44.

Pinto Prades JL. Is the person trade-off a valid method for allocating health care resources? Health Economics 1997;6:71-81

Raad voor de Volksgezondheid en Zorg. Waardebepaling van geneesmiddelen als beleidsinstrument. Advies aan de minister van VWS, mei 1997 (in Dutch) (Council for Health and Social Service. Evaluation of pharmaceuticals as a policy tool. Advice to the Dutch minister of Health, Welfare and Sport, May 1997). Zoetermeer, the Netherlands: Raad voor de Volksgezondheid en Zorg, 1997.

Ruwaard D, Kramers PGN (ed). Volksgezondheid Toekomst Verkenning 1997: De som der delen. (in Dutch) (Public Health Status and Forecast - report for 1997; available in English in 1998) Den Haag, the Netherlands: Elsevier/De Tijdstroom, 1997.

Ubel PA, Loewenstein G, Scanlon D, Kamlet M. Individual utilities are inconsistent with rationing choices: a partial explanation of why Oregon's cost-effectiveness failed. Med Decis Making 1996;16:108-16.

Verbrugge LM, Lepkowski JM, Imanaka Y. Comorbidity and its impact on disability. Milbank Q 1989;67(3-4),450-84.

Wetenschappelijke Raad voor het Regeringsbeleid. Volksgezondheidszorg. Rapporten aan de Regering nr. 52 (in Dutch) (Scientific Council for Government Policy. Public Health Care. Reports for the Government no. 52). Den Haag, the Netherlands: SDU, 1997

World Bank. World Development Report 1993: Investing in health - world development indicators. Oxford: Oxford University Press, 1993.

APPENDIX

Table A.1

Diagnostic groups, disease stages and EuroQol 5D+ classifications; incl. disability weights with 95% confidence intervals

Diagnostic group	Disease stages		EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
0. Terminal illness	0.5	end-stage disease otherwise unspecified	333332	0.5 interpolation (common core), chronic	0.07 (0.039;0.100)
1. Digestive tract infections	1.1P	digestive tract infection, uncomplicated course (duration 2 weeks)	112211	1.1P interpolation, annual profile	0.99 (0.991;0.999)
	1.2P	digestive tract infection, complicated course (duration 2-4 weeks)	323311	1.2P interpolation, annual profile	0.97 (0.961;0.982)
2. Meningitis	2.1	acute bacterial meningitis	333322	2.1 not valued	-
	-	deafness	see later		
	2.2	permanent locomotor impairment after bacterial meningitis	212111	2.2 interpolation, chronic	0.83 (0.702;0.964)
	2.3	permanent cognitive impairment after bacterial meningitis	112112	2.3 interpolation, chronic	0.75 (0.616;0.881)
	2.4	permanent locomotor and cognitive impairment after bacterial meningitis	213123	2.4 interpolation, chronic	0.24 (0.139;0.348)
3. Sepsis	3.1	septicaemia	333333	3.1 not valued	-

Diagnostic group	Disease stages		EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
4. STD - bacterial	4.1P	symptomatic acute gonorrhoea or Chlamydia trachomatis infection (duration 1 week)	111211	4.1P interpolation, annual profile	0.99 (0.981;0.995)
	4.2	late complications after gonorrhoeal or Chlamydia trachomatis infections	111221	4.2 interpolation, chronic	0.89 (0.801;0.968)
STD - viral	4.3	symptomatic non-fulminant acute hepatitis B infection	213211	4.3 interpolation, chronic	0.79 (0.707;0.862)
	4.4	chronic hepatitis B carriership without viral replication ('healthy carrier')'	111111 (50%), 111121 (50%)	4.4 interpolation, chronic	0.94 (0.913;0.966)
	4.5	chronic hepatitis B carriership with active viral replication	112221 (50%), 113321 (50%)	4.5 interpolation, chronic	0.64 (0.478;0.792)
	4.6	compensated liver cirrhosis	112221	4.6 interpolation, chronic	0.69 (0.546;0.834)
	4.7	decompensated liver cirrhosis	123322	4.7 interpolation, chronic	0.16 (0.040;0.273)
5. HIV/AIDS	5.1	seropositive	111121	5.1 interpolation, chronic	0.80 (0.696;0.897)
	5.2	AIDS-related complex	112121	5.2 interpolation, chronic	0.69 (0.496;0.891)
	5.3	AIDS - first stage	222221	5.4 interpolation, chronic	0.44 (0.324;0.556)
	5.4	AIDS - terminal	323222	5.4 not valued	-
6. Cancer of the oesophagus	6.1	stage of diagnosis and primary therapy	111221 (50%), 112331 (50%)	6.1 interpolation, chronic	0.44 (0.311;0.576)
	6.2	state after intentionally curative primary therapy	112221	6.2 interpolation, chronic	0.63 (0.576;0.691)
	6.3	irradically removed or disseminated carcinoma	112231 (50%), 113331 (50%)	6.3 interpolation, chronic	0.10 (0.069;0.134)
	6.4	preterminal stage	222231 (50%), 233332 (50%)	6.4 not valued	-
	6.5	terminal	333332	6.5 not valued	-
7. Cancer of the stomach	7.1	stage of diagnosis and primary therapy	111221(90%), 222331 (10%)	7.1 interpolation, chronic	0.47 (0.295;0.638)
	7.2	state after intentionally curative primary therapy	111221 (80%), 122231 (20%)	7.2 interpolation, chronic	0.62 (0.487;0.749)
	7.3	irradically removed or disseminated carcinoma	112231 (80%), 222331 (20%)	7.3 interpolation, chronic	0.27 (0.144;0.386)
	7.4	preterminal stage	222231 (80%), 222332 (20%)	7.4 not valued	-
	7.5	terminal	333332	7.5 not valued	-

Diagnostic group	Disease stages		EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
8. Colorectal cancer	8.1	stage of diagnosis and primary therapy	112231 (90%), 222231 (10%)	8.1 interpolation, chronic	0.57 (0.432;0.701)
	8.2	state after intentionally curative primary therapy	111121 (80%), 112221 (20%)	8.2 interpolation, chronic	0.80 (0.737;0.853)
	8.3	irradically removed or disseminated carcinoma	112231 (80%), 222331 (20%)	8.3 indicator condition	0.17 (0.129;0.210)
	8.4	preterminal stage	222231 (70%), 222332 (30%)	8.4 not valued	-
	8.5	terminal	333332	8.5 not valued	-
9. Lung cancer	9.1	stage of diagnosis and primary therapy for operable non-small cell lung cancer	112221 (60%), 123231 (40%)	9.1 interpolation, chronic	0.56 (0.417;0.692)
	9.2	stage of diagnosis and primary therapy for unoperable non-small cell lung cancer	123231 (50%), 223231 (50%)	9.2 interpolation, chronic	0.24 (0.157;0.313)
	9.3	non-small cell lung cancer, clinically diseasefree after primary therapy	112221	9.3 interpolation, chronic	0.53 (0.340;0.716)
	9.4	Disseminated non-small cell lung cancer	223332	9.4 interpolation, chronic	0.09 (0.056;0.124)
	9.5	Terminal	333332	9.5 not valued	
	9.6	Stage of diagnosis and chemotherapy for small-cell lung cancer	122221 (50%), 123231 (50%)	9.6 interpolation, chronic	0.32 (0.229;0.414)
	9.7	small-cell lung cancer, clinically 'in remission'	111121 (50%), 122231 (50%)	9.7 interpolation, chronic	0.46 (0.317;0.609)
	9.8	small-cell lung cancer, relapse/terminal	333332, 333333	9.8 not valued	-

Diagnostic group	Disease stages		EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
10 Breast cancer	10.1	diagnostic phase and primary therapy for non-invasive breast cancer or tumour < 2 cm	111221	10.1 interpolation, chronic	0.74 (0.648;0.828)
	10.2	diagnostic phase and primary therapy for breast tumour 2-5 cm. and/or local lymph node dissemination	112321	10.2 interpolation, chronic	0.31 (0.264;0.362)
	10.3	diagnostic phase and primary therapy for locally advanced breast cancer (tumour > 5 cm)	113331	10.3 interpolation, chronic	0.19 (0.137;0.236)
	10.4	clinically disease-free after the first year	111221	10.4 indicator condition	0.74 (0.663;0.817)
	10.5	disseminated	212331	10.5 interpolation, chronic	0.21 (0.163;0.260)
	10.6	terminal	323332	10.6 not valued	-
11. Prostate cancer	11.1	accidentally detected localised prostate cancer, follow-up without active intervention ('watchful waiting')	111121	11.1 interpolation, chronic	0.80 (0.736;0.861)
	11.2	diagnostic phase and primary therapy for localised prostate cancer	112221	11.2 interpolation, chronic	0.73 (0.647;0.803)
	11.3	clinically disease-free after primary therapy	111211 (50%), 111221 (50%)	11.3 interpolation, chronic	0.82 (0.743;0.904)
	11.4	disseminated	212221	11.4 interpolation, chronic	0.36 (0.191;0.526)
	11.5	hormone-refractory, terminal	323332	11.5 not valued	-

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
12. NHL	12.1 Non Hodgkin lymphoma of low-grade malignancy, dissemination stage I or II	111121 (50%), 111111 (50%)	12.1 interpolation, chronic	0.81 (0.731;0.885)
	12.2 Non Hodgkin lymphoma of low malignancy grade, dissemination grade III-IV)	111221 (80%), 112331 (20%)	12.2 interpolation, chronic	0.39 (0.275;0.504)
	12.3 Non Hodgkin lymphoma of intermediate/high malignancy grade, dissemination stage I	111121 (80%), 112221 (20%)	12.3 interpolation, chronic	0.45 (0.330;0.563)
	12.4 Non Hodgkin lymphoma of intermediary/high grade malignancy, dissemination stage II, III of IV	123331	12.4 interpolation, chronic	0.25 (0.168;0.338)
	12.5 terminal	233331	12.5 not valued	-
			Dissemination: I - one lymph node station II - two or more lymph node stations at the same side of the diaphragm III - lymph node stations at both sides of the diaphragm IV - disseminated disease in one or more organs and/or bone marrow	
13 Diabetes mellitus	13.1 uncomplicated	111111 (90%), 112221 (10%)	13.1 indicator condition	0.93 (0.906;0.953)
	13.2 with neuropathy	111111 (75%), 222221 (20%), 222331 (5%)	13.2 interpolation, chronic	0.81 (0.745;0.872)
	13.3 with nephropathy	112121 (80%), 113231 (20%)	13.3 interpolation, chronic	0.71 (0.620;0.799)
	- with other complications	see there		

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
14 Dementia	14.1 mild (only significant impairment of daily activities)	112112	14.1 interpolation, chronic	0.73 (0.582;0.871)
	14.2 moderate (independent living living is not possible without limited supervision)	123122	14.2 interpolation, chronic	0.37 (0.144;0.586)
	14.3 severe (permanent supervision required)	233123 (50%), 333133 (50%)	14.3 indicator condition	0.06 (0.046;0.073)
15 Schizophrenia	15.1 one psychotic episode, no permanent impairments	112111	15.1 interpolation, chronic	0.79 (0.649;0.930)
	15.2 several psychotic episodes, some permanent impairments	222122	15.2 interpolation, chronic	0.29 (0.212;0.364)
	15.3 several psychotic episodes, obvious permanent impairments	222223	15.3 interpolation, chronic	0.19 (0.099;0.281)
	15.4 several psychotic episodes, severe and increasing permanent impairments	233333	15.4 indicator condition	0.02 (0.016;0.023)
16 Depression	16.1 mild	112121	16.1 indicator condition	0.86 (0.806;0.914)
	16.2 moderate	122122	16.2 interpolation, chronic	0.65 (0.575;0.728)
	16.3 severe	223232	16.3 interpolation, chronic	0.24 (0.029;0.444)
	16.4 with psychosis, i.e. with delusions and/or hallucinations	223233	16.4 interpolation, chronic	0.17 (0.084;0.252)
17 Mental disorder	17.1 mild mental handicap (IQ=50-69)	112113	17.1 interpolation, chronic	0.71 (0.504;0.909)
	17.2 moderate mental handicap (IQ=35-49)	123113	17.2 interpolation, chronic	0.57 (0.482;0.647)
	17.3 severe mental handicap (IQ=20-34)	133113	17.3 interpolation, chronic	0.18 (0.033;0.327)
	17.4 extreme mental handicap (IQ<20)	233113	17.4 interpolation, chronic	0.24 (0.020;0.466)
	17.5 mental retardation (IQ = 70-84)	111112	17.5 interpolation, chronic	0.91 (0.863;0.960)

Diagnostic group	Disease stages		EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
18 Drug-related disorders, esp. alcohol	18.1	problem drinking (i.e., some physical, psychological or social problems caused by excessive alcohol intake)	112121	18.1 interpolation, chronic	0.89 (0.846;0.936)
	18.2	manifest alcoholism (severe social problems caused by excessive alcohol intake)	113221	18.2 interpolation, chronic	0.45 (0.271;0.635)
	18.3	psycho-organic disorder (delirium) caused by excessive alcohol intake	233233	18.3 interpolation, chronic	0.17 (0.015;0.314)
	-	liver disease (cirrhosis)	see there		
19 Psychic disorders in children and adolescents	19.1	autism (i.e., qualitative deficits in social interactions and communication	113123	19.1 interpolation, chronic	0.45 (0.290;0.606)
	19.2	mild behavioural disorder (hyperactivity)	111111	19.2 interpolation, chronic	0.98 (0.966;0.984)
	19.3	moderate to severe behavioural disorder (hyperactivity)	112111 (75%), 113111 (25%)	19.3 interpolation, chronic	0.85 (0.774;0.932)
	19.4	eating disorders (anorexia nervosa or bulimia nervosa)	111121 (80%), 112221 (20%)	19.4 interpolation, chronic [*children* with....]	0.72 (0.607;0.829)
20 M. Parkinson	20.1	initial stage (initially unilateral, later bilateral tremors and rigidity; slowness, impaired swallowing and speech; disturbance of equilibrium; patient are able to function indepedently)	212121	20.1 interpolation, chronic	0.52 (0.360;0.683)
	20.2	intermediate stage (swallowing and speech severely impaired; autonomic nervous system disturbances; patients are ADL-dependent, but are able to move without help)	223222	20.2 interpolation, chronic	0.21 (0.141;0.279)
	20.3	end-stage (wheelchair and bed patient, severely handicapped)	333232	20.3 interpolation, chronic	0.08 (0.051;0.102)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
21 Multiple sclerosis	21.1 'relapsing-remitting' phase	111111 (50%), 212121 (50%)	21.1 interpolation (common core), chronic	0.67 (0.616;0.722)
	21.2 progressive phase	222111 (50%), 333221 (50%)	21.2 interpolation, chronic	0.33 (0.232;0.434)
22 Vision disorders	22.1 mild (i.e., some difficulty reading small newspaper print, no difficulty recognizing faces at 4m. distance)	111111	22.1 interpolation, chronic	0.98 (0.967;0.999)
	22.2 moderate (i.e., great difficulty reading small newspaper print, some difficulty recognizing faces at 4m. distance)	112121	22.2 interpolation, chronic	0.83 (0.722;0.927)
	22.3 severe (i.e. unable to read small newspaper print, great difficulty or unable to recognize faces at 4m. distance)	123121	22.3 indicator condition	0.57 (0.479;0.661)
23 Hearing disorders in childhood	23.1 mild to moderate congenital or early required	112111	23.1 interpolation, chronic	0.89 (0.832;0.944)
	23.2 severe congenital or early acquired	113111	23.2 interpolation, chronic	0.77 (0.669;0.877)
24 Hearing disorders in elderly & 25 noise-related (acquired after early youth)	24.1 mild (i.e., some difficulty understanding or actively participating in a conversation with one or more persons)	111111	24.1 interpolation, chronic	0.96 (0.929;0.984)
	24.2 moderate (i.e., some difficulty to understand or participate in a conversation with one person but great difficulties with conversations with more than one person)	112121	24.2 interpolation, chronic	0.88 (0.844;0.906)
	24.3 severe hearing disorder in elderly (i.e., great difficulty or unable to understand or participate in a conversation with one other person)	113121	24.3 interpolation (common core), chronic	0.63 (0.594;0.659)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
26 Coronary heart disease (excl. heart failure)	26.1 mild stable angina pectoris (NYHA 1-2)	111121	26.1 indicator condition	0.92 (0.893;0.947)
	26.2 severe stable angina pectoris (NYHA 3)	212321	26.2 interpolation chronic	0.43 (0.281;0.582)
	26.3 instable angina pectoris	323231	26.3 not valued	-
	26.4 acute infarction	333332 (80%), ??? (20%)	26.4 not valued	-
27 Stroke	27.1 mild permanent impairments	112221	27.1 interpolation, chronic	0.64 (0.515;0.761)
	27.2 moderate permanent impairments	222222	27.2 indicator condition	0.37 (0.282;0.457)
	27.3 severe permanent impairments	233323	27.3 interpolation, chronic	0.08 (0.006;0.147)
28 Asthma / COPD	28.1 mild to moderate (symptom-free with or without maintenance therapy)	111111 (75%), 112111 (20%), 112121 (5%)	28.1 indicator condition	0.97 (0.949;0.990)
	28.2 severe asthma (not symptom-free despite maintenance medication)	112211 (75%), 113221 (25%)	28.2 interpolation, chronic	0.64 (0.573;0.714)
	28.3 mild-moderate COPD	112211	28.3 interpolation, chronic	0.83 (0.729;0.924)
	28.4 severe COPD	212221 (50%), 223231 (50%)	28.4 interpolation, chronic	0.47 (0.317;0.619)
29 Common cold	29.1P acute nasopharyngitis (duration 1 week)	111111 (85%), 112211 (15%)	29.1P interpolation, annual profile	1.00 (0.993;1.000)
	29.2P acute sinusitis (duration 2 weeks)	111211 (90%), 112211 (10%)	29.2P interpolation, annual profile	0.98 (0.968;0.991)
	29.3P acute tonsillitis (duration 2 weeks)	111211 (80%), 112211 (20%)	29.3P interpolation, annual profile	0.99 (0.991;0.999)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
30 Pneumonia, acute bronchitis and bronchiolitis	30.1P pneumonia (duration 2 weeks)	222222 (80%), 333332 (20%)	30.1P interpolation, annual profile	0.90 (0.809;0.984)
	30.2P acute bronchitis (duration 2 weeks, 1 episode a year)	112211 (50%), 113311 (50%)	30.2P interpolation, annual profile	0.99 (0.973;0.997)
	30.3P acute bronchitis (duration 2 weeks, more episodes a year)	112211 (50%), 113311 (50%)	30.3P interpolation, annual profile	0.96 (0.947;0.976)
	30.4 permanent stage after moderate to severe bronchiolitis	111211 (30%) / 112211 (10%)	30.4 interpolation, chronic [*children* with ...]	0.82 (0.692;0.951)
31 Influenza	31.1P influenza (duration 2 weeks)	111211 70%, 112211 20%, 333312 10%	31.1P interpolation (common core), annual profile	0.99 (0.980;0.992)
	31.1K influenza (duration 2 weeks)	111211 70%, 112211 20%, 333312 10%	31.1K interpolation (common core), short duration	0.84 (0.786;0.896)
32 Peptic ulcer disease	32.1P active gastric or duodenal peptic ulcer (duration 1 month)	111111 (20%), 111211 (60%), 112211 (10%), 112221 (10%)	32.1P interpolation, annual profile	0.98 (0.968;0.991)
33 Inflammatory bowel disease	33.1 inflammatory bowel disease, active exacerbation	111221 (40%), 112221 (40%), 223221 (20%)	33.1 interpolation, chronic	0.60 (0.505;0.695)
	33.2 inflammatory bowel disease, in remission	111111 (80%), 111221 (20%)	33.2 interpolation, chronic	0.82 (0.722;0.925)
34 Dental disease	34.1 dental caries	111111 (80%), 111221 (20%)	34.1 interpolation, chronic	0.99 (0.991;0.999)
	34.2 periodontal disease (gingivitis)	111111 (100%)	34.2 indicator condition	1.00 (0.999;1.000)
	34.3 periodontal disease (pockets > 6 mm. deep)	111111 (90%), 111211 (10%)	34.3 interpolation, chronic	0.99 (0.995;0.991)
	34.4 edentulism	111111 (75%), 111221 (25%)	34.4 interpolation, chronic	0.95 (0.916;0.981)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
35 Acute urinary tract infections	35.1P acute pyelitis / pyelonephritis (duration 2 weeks)	112221 (70%), 333321 (30%)	35.1P interpolation, annual profile	0.99 (0.976;0.996)
	35.2P acute urethritis (not STD) (duration 1 week)	111211	35.2P interpolation, annual profile	0.99 (0.977;0.999)
	35.3P acute cystitis (duration 1 week)	111211	35.3P interpolation, annual profile	0.99 (0.961;1.000)
36 Constitutional eczema	36.1 infant	112221	36.1 not valued	-
	36.2P 2 episodes of active constitutional eczema per year, of a duration of 6 weeks each	112211	36.2P interpolation, annual profile (2 times 6 weeks)	0.93 (0.874;0.993)
37 Contact eczema	see constitutional eczema			
38 Reumatoid arthritis	38.1 mild	122211	38.1 interpolation, chronic	0.79 (0.697;0.873)
	38.2 moderate	222221	38.2 interpolation, chronic	0.63 (0.485;0.781)
	38.3 severe	222331 (50%), 333331 (50%)	38.3 indicator condition	0.06 (0.039;0.080)
39 Osteoarthritis	39.1 grade 2 (radiological), hip or knee	111111 (70%), 211211 (10%), 212211 (10%), 222311 (10%)	39.1 interpolation, chronic	0.86 (0.776;0.940)
	39.2 grade 3-4 (radiological), hip or knee	111111 (20%) 222211 (60%), 222311 (10%), 333321 (5%), 233321 (5%)	39.2 interpolation, chronic	0.58 (0.361;0.796)
40 Osteoporosis	40.1 2 SD below normal (WHO-definition)	111111	40.1 not valued (risk factor)	-
41 Neural tube defects	41.1 young adults with high level spina bifida aperta (L2 or higher)	322211 (60%), 333212 (40%)	41.1 interpolation, chronic	0.32 (0.187;0.453)
	41.2 young adults with medium level spina bifida aperta (L3 to L5)	212211 (75%) 322212 (25%)	41.2 interpolation, chronic	0.50 (0.474;0.525)
	41.3 young adults with a low spina bifida aperta (sacral)	112211	41.3 interpolation, chronic	0.84 (0.747;0.926)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
42 Congenital heart disease	42.1 young adult in permanent stage after intentionally curative operation for congenital atrial or ventricular septal defect	111111	42.1 interpolation, chronic	0.97 (0.952;0.991)
	42.2 child/adolescent in permanent stage after intentionally curative operation for Fallot's tetralogy or transposition of the great arteries	112221	42.2 interpolation, chronic	0.80 (0.687;0.909)
	42.3 young adult in permanent stage after intentionally curative operation for Fallot's tetralogy or transposition of the great arteries	112211	42.3 interpolation, chronic	0.89 (0.846;0.930)
	42.4 child in permanent stage after intentionally curative operation for pulmonary stenosis	111111	42.4 interpolation, chronic	0.98 (0.959;0.997)
	42.5 young adult in permanent stage after intentionally curative operation for pulmonary stenosis	112211	42.5 interpolation, chronic	0.84 (0.687;0.999)
	42.6 child/adolescent in permanent stage with complex not curatively operable congenital heart disease	113321	42.6 interpolation, chronic	0.28 (0.186;0.380)
43 Down's syndrome	43.1 child, age below 10 with Down's syndrome, with other congenital anomalies	333213	43.1 interpolation, chronic	0.31 (0.103;0.509)
	43.2 child, age below 10, with Down's syndrome, without other congenital anomalies	122113	43.2 interpolation, chronic	0.49 (0.425;0.554)
	43.3 patient (10 - 40 years) with Down's syndrome	122113	43.3 interpolation, chronic	0.65 (0.420;0.873)
	43.4 adult, over 40 years of age, with Down's syndrome	133223	43.4 interpolation, chronic	0.35 (0.138;0.565)
44 Premature birth (excl. congenital anomalies)	44.1 children with permanent impairments 5 years after premature birth (< 32 weeks)	222122	44.1 interpolation, chronic	0.52 (0.466;0.574)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
45 Health problems excl. congenital anomalies in maturely-born children	45.1 children with permanent impairments after dysmature birth ('small for gestational age', birth weight < 5th percentile)	212122	45.1 interpolation, chronic	0.65 (0.484;0.815)
	45.2 children with permanent impairments after asphyxia (APGAR < 7 after 5 minutes)	222122	45.2 indicator condition	0.51 (0.406;0.614)
	45.3 children with permanent impairments after perinatal bacterial infection	222112	45.3 interpolation, chronic	0.64 (0.437;0.843)
	45.4 children with permanent impairments after perinatal viral infection	111112 (60%), 222123 (40%)	45.4 interpolation, chronic	0.54 (0.444;0.643)
46 Complications of multiple gestation	not valued (see 44 and 45)			
47/48. Accidents & Injuries	47.1 permanent impairments after mild skull/brain injury	111212 (60%), 111223 (40%)	47.1 interpolation, chronic	0.63 (0.487;0.763)
	47.2 permanent impairments after moderately severe skull/brain injury	222222 (50%),222223 (50%)	47.2 interpolation, chronic	0.27 (0.188;0.343)
	47.3 permanent impairments after severe skull/brain injury	222223 (75%), 333333 (25%)	47.3 interpolation, chronic	0.26 (0.083;0.433)
	47.4 paraplegia, stable stage	222111 (85%), 332221 (15%)	47.4 indicator condition	0.43 (0.349;0.511)
	47.5 tetraplegia, stable stage	332111 (70%), 333221 (30%)	47.5 interpolation, chronic	0.16 (0.063;0.257)
	47.6 permanent impairments after fracture of arm or shoulder	122111	47.6 interpolation, chronic	0.94 (0.906;0.964)
	47.7 permanent impairments after fracture of leg or hip	222111	47.7 interpolation, chronic	0.87 (0.793;0.947)
	47.8 permanent impairment after luxation or distorsion of ankle or foot	212211	47.8 interpolation, chronic	0.97 (0.950;0.986)
	47.9 permanent impairments after burns	112121	47.9 interpolation, chronic	0.86 (0.771;0.957)
49. Tuberculosis	49.1 tuberculosis of the lung	112211 (40%), 222221 (60%),	49.1 interpolation, chronic	0.71 (0.594;0.819)
	49.2 'remnant TB'	112211 (10%)	49.2 interpolation, chronic	0.84 (0.760;0.919)
	49.3 extrapulmonary tuberculosis	112211 (80%), 223321 (20%)	49.3 interpolation, chronic	0.70 (0.538;0.864)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
50 Skin cancer (incl. melanoma)	50.1 basal cell carcinoma	111111	50.1 interpolation, chronic	0.95 (0.909;0.980)
	50.2 squamous cell skin cancer, undisseminated	111111 (80%), 111121 (20%)	50.2 interpolation, chronic	0.93 (0.881;0.975)
	50.3 squamous cell skin carcinoma with lymph node dissemination	111221	50.3 interpolation, chronic	0.60 (0.449;0.744)
	50.4 malignant melanoma I, no evidence of disseminationmelanoom I	111121	50.4 interpolation, chronic	0.81 (0.730;0.883)
	50.5 malignant melanoma II, lymph node dissemination, no distant dissemination	111121 (60%), 111131 (40%)	50.5 interpolation, chronic	0.57 (0.365;0.764)
	50.6 malignant melanoma III, disseminated	111121 (60%), 111231 (40%)	50.6 interpolation, chronic	0.19 (0.106;0.280)
	50.7 terminal	223332	50.7 not valued	-
51. Anxiety disorders	51.1 mild to moderate panic disorder	112121	51.1 interpolation, chronic	0.84 (0.765;0.914)
	51.2 severe panic disorderis	113131	51.2 interpolation, chronic	0.31 (0.226;0.393)
	51.3 mild to moderate agoraphobia	112121	51.3 interpolation, chronic	0.89 (0.838;0.934)
	51.4 severe agoraphobia	113132	51.4 interpolation, chronic	0.45 (0.301;0.588)
	51.5 mild to moderate singular phobia	111121	51.5 interpolation, chronic	0.88 (0.860;0.889)
	51.6 severe singular phobia	112131	51.6 interpolation, chronic	0.58 (0.379;0.787)
	51.7 mild to moderate social phobia	112121	51.7 interpolation, chronic	0.83 (0.765;0.901)
	51.8 severe social phobia	113131	51.8 interpolation, chronic	0.41 (0.212;0.611)
	51.9 mild to moderate obsessive-compulsive disorder	112122	51.9 interpolation, chronic	0.76 (0.679;0.834)
	51.10 severe obsessive-compulsive disorder	122133	51.10 interpolation, chronic	0.44 (0.259;0.620)
	51.11 mild to moderate posttraumatic stress disorder	112121	51.11 interpolation (common core), chronic	0.87 (0.847;0.891)
	51.12 severe posttraumatic stressdisorder	112132	51.12 interpolation, chronic	0.49 (0.343;0.629)
	51.13 mild to moderate diffuse anxiety disorder	112121	51.13 interpolation, chronic	0.83 (0.792;0.871)
	51.14 severe diffuse anxiety disorder	112232	51.14 interpolation, chronic	0.40 (0.280;0.523)
52 Epilepsy	52.1 epilepsy	112111	52.1 interpolation, chronic	0.89 (0.838;0.948)

Diagnostic group	Disease stages	EQ 5D+ classification	Remarks[a]	Diability weight (95% C.I.)
53 Heart failure	53.1 mild (NYHA 1 - 2)	111211	53.1 interpolation, chronic	0.94 (0.921;0.962)
	53.2 moderate (NYHA 3)	222211	53.2 interpolation, chronic	0.65 (0.481;0.815)
	53.3 severe (NYHA 4)	223321	53.3 interpolation (common core), chronic	0.35 (0.296;0.405)
54 Low back pain	54.1 low back pain	212211	54.1 indicator condition	0.94 (0.916;0.963)
55. Hip fracture	55.1 during rehabilitation	222211	55.1 interpolation, chronic	0.81 (0.688;0.935)
	47.7 after 1 year	see there		
56. ADL-limitations	56.1 none to mild ADL limitations in elderly	111111	56.1 interpolation, chronic	0.99 (0.988;0.994)
	56.2 moderate to severe ADL limitations in elderly	222111	56.2 indicator condition	0.89 (0.836;0.944)
	56.3 elderly with extreme ADL limitations or complete ADL dependence	333111	56.3 interpolation, chronic	0.35 (0.282;0.411)

a. Number of observations: indicator conditions n=34; common core interpolation n=38, interpolation n=6