Cross-national comparability of burden of disease estimates: the European Disability Weights Project

Marie-Louise Essink-Bot,[1] Joaquin Pereira,[2] Claire Packer,[3] Michael Schwarzinger,[4] Kristina Burström,[5] & the European Disability Weights Group[6]

Objective To investigate the sources of cross-national variation in disability-adjusted life-years (DALYs) in the European Disability Weights Project.

Methods Disability weights for 15 disease stages were derived empirically in five countries by means of a standardized procedure and the cross-national differences in visual analogue scale (VAS) scores were analysed. For each country the burden of dementia in women, used as an illustrative example, was estimated in DALYs. An analysis was performed of the relative effects of cross-national variations in demography, epidemiology and disability weights on DALY estimates.

Findings Cross-national comparison of VAS scores showed almost identical ranking orders. After standardization for population size and age structure of the populations, the DALY rates per 100 000 women ranged from 1050 in France to 1404 in the Netherlands. Because of uncertainties in the epidemiological data, the extent to which these differences reflected true variation between countries was difficult to estimate. The use of European rather than country-specific disability weights did not lead to a significant change in the burden of disease estimates for dementia.

Conclusions Sound epidemiological data are the first requirement for burden of disease estimation and relevant between-countries comparisons. DALY estimates for dementia were relatively insensitive to differences in disability weights between European countries.

Keywords Disability evaluation; Cost of illness; Dementia/epidemiology; Women; Data collection/methods; Analysis of variance; Comparative study; Europe; Denmark; United Kingdom; France; Netherlands; Spain; Sweden (*source: MeSH, NLM*).

Mots clés Evaluation incapacité; Coût maladie; Démence/épidémiologie; Femmes; Collecte données/méthodes; Analyse variance; Etude comparative; Europe; Danemark; Royaume-Uni; France; Pays-Bas; Espagne; Suède (*source: MeSH, INSERM*).

Palabras clave Evaluación de la incapacidad; Costo de la enfermedad; Demencia/epidemiología; Mujeres; Recolección de datos/métodos; Análisis de varianza; Estudio comparativo; Europa; Dinamarca; Reino Unido; Francia; Países Bajos; España; Suecia (*fuente: DeCS, BIREME*).

Bulletin of the World Health Organization 2002;80:644-652.

Voir page 650 le résumé en français. En la página 650 figura un resumen en español.

Introduction

The value of summary measures of population health as a tool for health policy and planning purposes has been increasingly recognized (*1*). Summary measures such as the disability-adjusted life-year (DALY) combine mortality data with that on non-fatal health outcomes. In the DALY, as developed in the Global Burden of Disease (GBD) study, mortality and morbidity were included as years of life lost (YLL) and years lived with disability (YLD), respectively (*2*). Essential information for the estimation of DALYs include disease-specific epidemiological data and disability weights. The latter are used for weighting the years lived with a specific disease by the severity of the disability associated with it. The GBD study inspired the use of DALYs in comprehensive studies on disease burden at the national level (*3*–*5*). Both the valuation technique and the use of expert opinion in order to elicit disability weights in the GBD study have been questioned on methodological grounds (*6*–*8*). Additionally, the sensitivity of DALYs, defined by the relative contributions of true and error variation, is assumed to be low. Potential sources of true variation include differences in the size and structure of populations, real differences in disease epidemiology between populations or over time, and differences in disability weights. Error variation may originate from sampling and measurement error and from incomparability in available epidemiological data and disability weights. The detection of true variation is

[1] Department of Public Health, Erasmus Medical Center Rotterdam, PO Box 1738, 3000 DR Rotterdam, Netherlands (email: essink@mgz.fgg.eur.nl). Correspondence should be addressed to this author.

[2] National School of Public Health, Carlos III Institute of Health, Madrid, Spain.

[3] Department of Public Health and Epidemiology, University of Birmingham, Birmingham, England.

[4] Public Health, Henri Mondor Hospital, AP-HP, Paris, France.

[5] Department of Public Health Sciences, Division of Social Medicine, Karolinska Institute, Stockholm, Sweden.

[6] Finn Kamper-Jørgensen, Ulla Christensen, Kim Moesgaard Iburg (Denmark); James Raftery, Claire Packer, Lisa Gold, Suzanne Robinson (from October 1999) (England); Isabelle Durand-Zaleski, Michael Schwarzinger (France); Louise Gunning-Schepers, Gouke Bonsel, Clara Moerman (from January 2000), Marlies Stouthard, Paul van der Maas (project coordinator), Marie-Louise Essink-Bot (Netherlands); Joaquin Pereira, Ana Baylin (until January 1999), Eduardo Fernandez Zincke (Spain); Finn Diderichsen, Kristina Burström, Rickard Ljung (Sweden).

Ref. No. **01-1421**

the focus of interest when estimating the burden of disease in DALYs. However, error may limit the power to detect true differences between populations (*9*).

The cross-national stability of disability weights and the relative effects of three sources of variability on DALY estimates were investigated in the European Disability Weights Project. In this paper we present the derivation of country-specific disability weights for 15 disease stages; analyse cross-national variation in visual analogue scale (VAS) scores for these disease stages; apply the disability weights in a burden-of-disease estimation of dementia in women in Denmark, England and Wales, France, the Netherlands, Spain, and Sweden; and explore the relative importance of cross-national differences in demography, epidemiology, and disability weights.

Methods

The European Disability Weights Project included participants from Denmark, England and Wales, France, the Netherlands, Spain, and Sweden, and ran for two years from March 1998.

The disability weight derivation consisted of the following: disease selection and staging; valuation; and analysis with special attention to country-specific effects.

The burden of disease estimation involved the following: collection of demographic data; collection of prevalence data for dementia; epidemiological modelling; and DALY estimation.

Dementia was chosen because there was evidence that good and comparable epidemiological data were available in the participating countries. Only the results for women are presented because of limited space.

Disease selection and staging

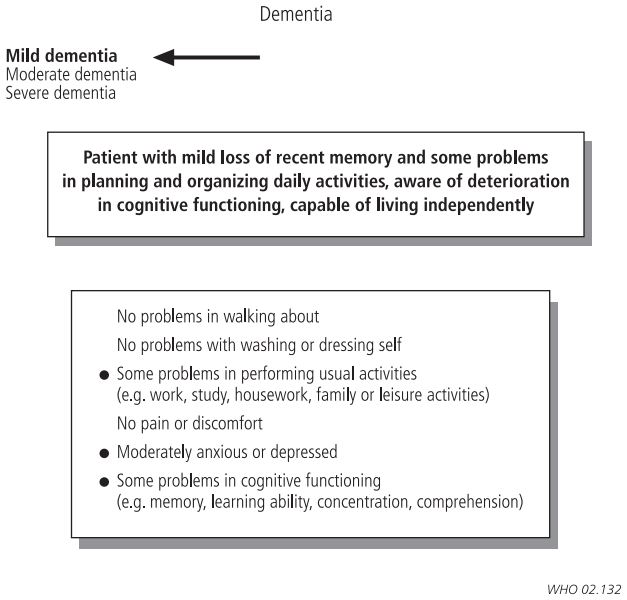
Thirteen diseases were selected which covered a range of severity and different dimensions of disability and were relevant for the European situation. Each was subdivided into homogeneous stages with respect to functional status, treatment, and prognosis. A textual and a standardized generic description of the associated functional health status, validated by clinical experts, was developed for each stage. The generic description used the EuroQol (EQ-5D) classification system of health status (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), with the addition of a sixth dimension of cognitive functioning (EQ-5D+C) (*10*, *11*). Fig. 1 provides an example of a description of a disease stage as presented to panellists in the valuation procedure.

Valuation procedure

The valuation procedure, modified from the GBD protocol and the Dutch Disability Weights Study (*12*, *13*), used a panel process to value 15 disease stages, one or two for each selected disease. Subsequently a postal interpolation procedure was used for the remaining stages of the diseases. Three valuation methods, preceded by a ranking procedure, were employed in the panel sessions: VAS method with 15 disease stages; time trade-off (TTO) with nine disease stages; and a newly developed variant of person trade-off (PTO) with nine disease stages.

In the VAS method, panellists located the health state descriptions on a scale with anchored end-points ("best" to "worst" imaginable health state) in order of preference. In

Fig. 1. **Example of a disease description as presented for valuation, including a disease label (dementia), the disease stage to be valuated (marked by an arrow), a textual description (in bold) and a generic description of the functional health status (EQ-5D+C; three severity levels per attribute; dot indicates the moderate level)**



TTO, panellists hypothetically traded off a number of healthy life years in order to avoid years lived in the health state being valued. We used the standard individual TTO with a time horizon of 10 years (*14*). In PTO, panellists were asked to prioritize between two preventive programmes A and B, where A saved one hundred healthy lives and B prevented a chronic disease state. Panellists determined the number of persons in B if they were indifferent between A and B.

The valuation protocol and disease descriptions, written in English, were translated into the languages of the participating countries and subjected to a formal process of forward-and-backward translation. Panel facilitators received standardized training. A total of 23 panel sessions with 232 panel members were held in five countries. Ten of the panels involved non-health professionals and 13 were made up of health professionals. Empirical disability weights were not derived in Denmark.

The results of the TTO and PTO valuations, including attention to valuations by health and non-health professionals, are to be published elsewhere. The complete valuation protocol is available from the authors on request.

Analysis of valuation data with special attention to country-specific effects

We used an analysis of variance (ANOVA) to evaluate differences in VAS scores across countries for each disease stage. After adjustment for multiple testing, involving the use of the Bonferoni equation, a P-value <0.003 was considered statistically significant. An analysis of the components of variance (generalizability study or G-study), involving the use of maximum likelihood estimation (*15*), was used to investigate whether between-country differences were systematic and to study the relative impact of various sources on the total variance of the valuations. Among the results of the G-study,

the G-coefficient represents the proportion of variance attributable to true differences that can be interpreted as an internal consistency reliability coefficient for the valuations.

VAS scores, which do not include an explicit trade-off and, in our case, did not include death as an anchor point, are not suitable for direct use as disability weights. A European transformation function based on pooled data, and five country-specific transformation functions, all with the general form proposed by Torrance (TTO = 1–(1–VAS)$^{\alpha}$), were estimated on the basis of panel VAS and TTO data for nine disease stages (*14*). Alpha coefficients in the transformation function were 0.51 for the pooled data (0.41 for Sweden, 0.45 for England and Wales, 0.46 for France, 0.54 for Spain, and 0.69 for the Netherlands). VAS scores from panel valuations and the interpolation were transformed into TTO values with the respective country-specific transformation functions or with the European transformation function in order to obtain pooled disability weights.

Collection of demographic data

We collected national population and total mortality data for 1996 in five-year age groups up to ⩾95 years from the national statistical offices of the respective countries: Statistics Denmark, the Office for National Statistics in England and Wales, the French National Institute of Statistics and Economic Studies, Statistics Netherlands, the Spanish National Institute of Statistics, and Statistics Sweden.

Collection of prevalence data for dementia

We used primary prevalence data from recent population surveys on dementia in the respective countries. Although the screening tests differed, all the surveys included the Mini-mental State Examination (*16–22*). Further details of the studies are given in Annex 1.

Epidemiological modelling

Because prevalence data covering the complete age range were not available in all countries, we modelled the missing prevalences of female dementia for age groups 55–59 years to ⩾95 years by means of a dementia prevalence model (*23*) (Annex 2). For each country a composite disability weight for dementia was calculated on the basis of the prevalence distribution across the mild, moderate, and severe stages of dementia and the country-specific disability weights. We used the observed relationship between prevalence and severity distribution as reported from each country-specific study in order to model mutually comparable country-specific severity distributions. For Denmark the pooled European disability weights were used.

DALY estimation

Years lived with disability were estimated from an incidence perspective by multiplying age-specific female dementia incidence by average duration and composite disability weight. The female age-specific incidence was modelled using an incidence–prevalence–mortality model (DisMod) from prevalence data, an estimate of the excess mortality risk in dementia, and an assumed remission of zero (*24*). Because age-specific and sex-specific relative risks of dementia mortality were not available for all the countries, these risks were calculated from the estimates of the mortality rate ratios available from the Rotterdam study (*20*). The Mantel–Haenszel combined estimates of the mortality rate ratios in this study were 2.1 for men (95% confidence interval (CI) = 1.5–2.9) and 2.3 for women (95% CI = 1.9–2.9). Finally, years lived with disability attributable to dementia, were estimated as: [incidence (from DisMod) × duration (from DisMod) × composite disability weight].

Years of life lost for dementia were estimated by multiplying female age-specific dementia mortality rates by female age-specific standard expected years of life lost and female population numbers. Standard expected YLL were derived from the standard life table West 26, modified (with a life expectancy at birth of 82.5 years for females) by using Australian estimates of the mean age at death per five-year age group (*25*). YLL estimates were not based on registered dementia mortality but on dementia mortality as modelled in DisMod. Thus dementia was regarded as a risk factor for mortality but not as a specific cause of mortality. Because of the uncertainty in the literature concerning age-weighting and age-discounting in DALY calculations (*26–28*), these procedures were not undertaken in the present study.

In order to illustrate the relative contribution of cross-national variations in demography and disability weights, we calculated dementia DALYs for women as follows: in absolute numbers, using country-specific population figures and country-specific disability weights; in DALY crude rates per 100 000 women and with country-specific disability weights; in standardized DALY rates per 100 000 women, using the European standard population (*29*) and country-specific disability weights; and in standardized DALY rates per 100 000 women, using the European standard population and pooled European disability weights.

Results

Valuations

Table 1 shows the VAS scores obtained in panel sessions for the 15 disease stages. Differences between countries appeared to be significant (P <0.003) for seven of the disease stages. Spearman rank correlation between rank orders in countries (two by two) yielded an average r_s of 0.97 (range = 0.94–0.99).

Variance component analysis indicated that "disease stage" explained 67% of variance (Table 2). Little or no variance was explained by "country", "panel nested within country", "individual nested within panel", "age" or "sex". Interactions between "disease stage" and "panel" and between "disease stage" and "country" accounted for small components of variance. This indicated that, in some countries and in some panels within countries, some disease stages were valued differently. Adding more variables or first-order interaction terms did not reduce the total proportion of unexplained variance (25%). The G-coefficient for the total international data, i.e. the proportion of variance attributable to differences in scores of the disease stages, was 0.67. The G-coefficients at the national level were 0.61 for France, 0.63 for England and Wales, 0.71 for Sweden, 0.74 for the Netherlands, and 0.76 for Spain. A total of 6% of the variance in the French data set was attributable to first-order interaction between "disease stage" and "panel", indicating that in France some disease stages were valued differently in some panels. In the other countries only 1% of the variance was attributable to this first-order interaction. "Panel member nested within panel" accounted for 4% of the variance in the Netherlands, 5% in England and

Table 1. **Differences across countries per disease stage in terms of VAS scores (0 = no disability, 1= maximum disability) for 15 disease stages obtained in panel sessions in England and Wales, France, the Netherlands, Spain, and Sweden**

Disease stage	Mean ± standard deviation:						ANOVA
	Pooled (*n* = 231)	England and Wales (*n* = 49)	France (*n* = 46)	Netherlands (*n* = 50)	Spain (*n* = 47)	Sweden (*n* = 39)	
Common cold	0.03 ± 0.04	0.04 ± 0.06	0.03 ± 0.02	0.03 ± 0.05	0.04 ± 0.04	0.03 ± 0.03	F = 0.844 *P* = 0.498
Vision disorder (mild/moderate)	0.17 ± 0.17	0.17 ± 0.14	0.29 ± 0.29	0.13 ± 0.06	0.15 ± 0.11	0.14 ± 0.08	F = 7.336 *P* = 0.000
Breast cancer (disease-free, no permanent sequelae)	0.32 ± 0.18	0.37 ± 0.22	0.29 ± 0.19	0.27 ± 0.11	0.39 ± 0.15	0.26 ± 0.14	F = 5.595 *P* = 0.000
Low back pain	0.33 ± 0.18	0.37 ± 0.20	0.34 ± 0.17	0.28 ± 0.15	0.32 ± 0.16	0.36 ± 0.20	F = 2.124 *P* = 0.079
Diabetes (difficult to control)	0.34 ± 0.16	0.40 ± 0.17	0.39 ± 0.14	0.31 ± 0.10	0.31 ± 0.14	0.37 ± 0.16	F = 3.615 *P* = 0.007
Asthma (severe)	0.46 ± 0.19	0.53 ± 0.20	0.43 ± 0.23	0.43 ± 0.14	0.43 ± 0.17	0.48 ± 0.15	F = 2.714 *P* = 0.031
Dementia (mild)	0.46 ± 0.21	0.50 ± 0.23	0.46 ± 0.23	0.46 ± 0.17	0.47 ± 0.19	0.41 ± 0.21	F = 0.965 *P* = 0.428
Colorectal cancer (diagnosis + primary therapy)	0.51 ± 0.20	0.53 ± 0.22	0.48 ± 0.21	0.47 ± 0.14	0.64 ± 0.15	0.41 ± 0.18	F = 10.318 *P* = 0.000
HIV (seropositive, minor)	0.55 ± 0.22	0.56 ± 0.21	0.60 ± 0.22	0.43 ± 0.19	0.65 ± 0.21	0.52 ± 0.20	F = 7.381 *P* = 0.000
Myocardial infarction	0.59 ± 0.20	0.64 ± 0.22	0.61 ± 0.22	0.49 ± 0.13	0.67 ± 0.16	0.52 ± 0.18	F = 7.760 *P* = 0.000
Angina pectoris (severe stable)	0.59 ± 0.16	0.65 ± 0.15	0.58 ± 0.18	0.54 ± 0.13	0.59 ± 0.14	0.56 ± 0.15	F = 3.645 *P* = 0.007
Stroke (moderate permanent impairments)	0.68 ± 0.16	0.69 ± 0.17	0.68 ± 0.17	0.64 ± 0.13	0.75 ± 0.13	0.64 ± 0.15	F = 3.460 *P* = 0.009
Depression (severe)	0.78 ± 0.17	0.74 ± 0.22	0.77 ± 0.21	0.82 ± 0.12	0.76 ± 0.13	0.83 ± 0.14	F = 2.568 *P* = 0.000
Quadriplegia	0.91 ± 0.10	0.89 ± 0.12	0.93 ± 0.07	0.89 ± 0.05	0.94 ± 0.06	0.88 ± 0.13	F = 2.859 *P* = 0.024
Final year of disease	0.91 ± 0.10	0.92 ± 0.09	0.96 ± 0.05	0.87 ± 0.07	0.95 ± 0.06	0.87 ± 0.12	F = 10.167 *P* = 0.000

Wales, and Spain, 6% in Sweden, and 2% in France. "Panel" accounted for 0% of the variance in all countries. Unexplained variance amounted to 18% in Spain, 21% in the Netherlands, 22% in Sweden, 31% in France, and 33% in England and Wales.

Burden of disease

Prevalence data were available for all countries only for the age groups 75–79 years and 80–84 years, and varied considerably between countries (Fig. 2). In general, the lowest prevalences were those of England and Wales, France, and Spain, and the highest were those of the Netherlands. After estimating the missing prevalences, we modelled the complete age-specific female dementia incidences for each country (Fig. 3).

High absolute numbers of DALYs occurred in England and Wales (539 106) and France (526 441), while the number in Denmark was much lower (55 958). These differences were largely determined by differences in population size: in mid-1996 the estimated female populations in Denmark, England and Wales, and France were 2 667 015, 26 452 800 and 29 892 763, respectively.

The calculation of DALY crude rates per 100 000 women indicated that the largest burden of dementia occurred in England and Wales (data not shown). Application to the standard European female population resulted in very similar distributions across the study countries, with a maximum burden of dementia attributed to the age group 70–74 years (Fig. 4). The highest standardized burden was observed in the Netherlands (1404), the lowest in France (1050). The standardized female dementia burden amounted to 1093 in Sweden, 1129 in Spain, 1235 in England and Wales, and 1253 in Denmark.

The use of European weights instead of country-specific weights resulted in only small changes in the standardized burden of dementia. For Spain and the Netherlands, the application of European disability weights for dementia decreased the total burden in DALYs by 2.2% and 5.9%, respectively. For Sweden, France, and England and Wales, the same procedure led to increases of 6.4%, 2.1%, and 1.0%, respectively, in the total burden.

Table 2. **Variance components of VAS scores for 15 disease stages in England and Wales, France, the Netherlands, Spain, and Sweden**

Source of variance	Variance component
Disease stages	0.0581 (67)[a]
Country	0.0006 (1)
Panel nested within country	0.0000 (0)
Panel member nested within panel nested within country	0.0035 (4)
Age	0.0003 (0)
Sex	0.0000 (0)
Disease stages × panel nested within country	0.0016 (2)
Disease stages × country	0.0013 (1)
Disease stages × age	–
Disease stages × sex	–
Unexplained	0.0215 (25)

[a] Figures in parentheses are percentages.

Fig. 2. **Observed prevalence of dementia among females in Denmark, England and Wales, France, the Netherlands, Spain, and Sweden**

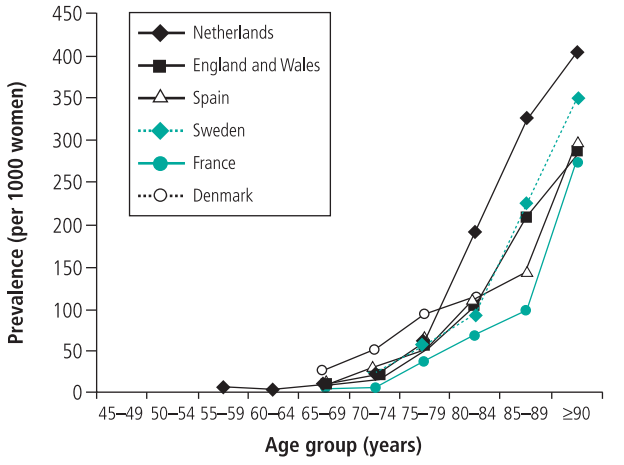


Fig. 3. **Modelled dementia incidence for females in Denmark, England and Wales, France, the Netherlands, Spain, and Sweden**

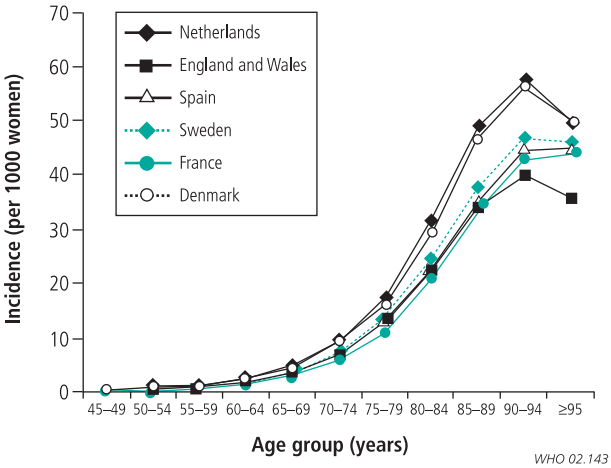
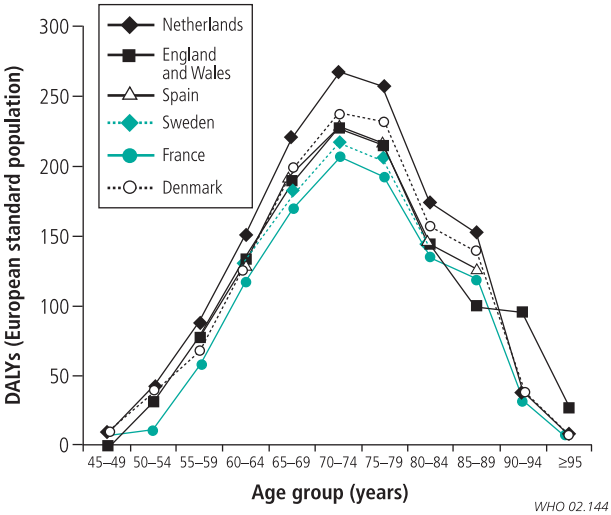


Fig. 4. **Burden of dementia in standardized DALY rates per 100 000 women in Denmark, England and Wales, France, the Netherlands, Spain and Sweden in 1996, based on country-specific dementia prevalence data and country-specific disability weights, and using the female European standard population**



Discussion

Visual analogue scale results

The analysis of cross-national differences between VAS scores obtained in panel sessions for the 15 disease stages showed an almost identical rank order of stages across countries. Real differences between the VAS scores of different health states explained the largest part of variance in the total data set. The observed cross-national differences did not appear to be systematic but were largely attributable to unexplained variance. The internal consistency of VAS scores ranged from 0.61 in France to 0.76 in Spain. The proportion of unexplained variance showed the reverse pattern, the highest error variances being in France and in England and Wales, and the lowest being in Spain. Generally, a norm of 0.70 or higher is recommended for internal consistency reliability (*30*). The differences between countries in the proportions of variance attributable to the various components may be regarded as an indicator that, despite all our efforts, cross-national standardization of the valuation protocol, including the selection of panel members, written instructions for the facilitators and careful translation procedures, was not achieved.

In the panel sessions, TTO and PTO values were also derived. We have compared only the VAS scores that theoretically provided the best opportunity for detection of a variance component attributable to differences between countries because the percentage of unexplained variance in the VAS data (25%) was lower than in the TTO and PTO data sets. The analysis of VAS scores can be further justified because our TTO and PTO methods were not applicable to six of the disease stages and because some panel members had problems with either TTO or PTO. VAS thus gave the most complete and consistent data set. It thus appears that VAS and subsequent transformation into TTO values is a usable method for determining disability weights on a larger scale.

Both GBD and the Dutch Disability Weights Study derived disability weights for diseases that were also evaluated in the European Disability Weights Project. However,

comparison is not straightforward because of differences in the descriptions of the stages that were evaluated, the valuation methods, and the scope of the valuations. The disability weights from the present study cannot replace the GBD weights because the European Disability Weights Project covers only a limited number of conditions: it aimed at studying cross-national differences in disability weights in Western countries and their effects on burden of disease estimates.

Burden of disease from dementia

The application of country-specific disability weights had a relatively small effect on the estimated burden of dementia and the results did not support the use of country-specific weights in this instance.

Detailed burden of disease estimation for dementia was a demanding epidemiological exercise. The extensive population data proved to be of limited applicability: national studies focused on different age groups and there was a lack of data on the excess mortality risk. For other diseases such as cancer, the general quality of epidemiological data may be better because there are long-standing national cancer registries in many countries, but even in this instance cross-national treatment differences may hamper comparability. For many other diseases, and in other parts of the world, the quality of data is probably much worse. In respect of dementia, extensive analysis and modelling were required before the DALY calculations were made. The analysis was based on the strong assumption that the course of dementia was similar across countries. This can be defended from a biological viewpoint, although cross-national differences in vascular risk profiles cannot be ruled out. Such an assumption causes difficulties in the interpretation of differences in estimates of disease burden across countries.

The results of the present study illustrate the sensitivity of DALY estimates to three sources of variation between countries. Demographic differences are a source of true variation. The precision of the YLL estimates in our study, i.e. the degree of freedom from error variance, largely depends on the values of the mortality risk ratios used to estimate mortality attributable to dementia. As far as the YLDs are concerned, the main sources of uncertainty are in the prevalence estimates, the disability weights and the severity distribution of dementia. For the European VAS scores, we found unexplained variance of 25%. In the Australian burden of disease study using iterative sampling techniques, an uncertainty analysis revealed an estimated relative standard error of 12% in the YLD estimates for dementia (*5*).

Because different coding practices make it impossible to compare registered dementia mortality internationally, YLL estimates were not based on registered dementia mortality but on dementia mortality as modelled in DisMod (*31*). This method of estimating YLL, considering dementia as a risk factor for mortality, enhanced the comparability of burden of disease estimates for dementia across countries. However, these DALY estimates are not comparable with estimates based on registered dementia mortality data or with DALY estimates for other diseases where mortality registry data are used.

We presented a cross-national comparison of DALY estimates for a single disease in order to show the sensitivity of DALY estimates to various sources of variance. However, the DALY approach was not originally developed to compare burdens of single diseases. Rather, it was intended to provide a common denominator for comparison of burdens of different diseases and to enable the estimation of an internally consistent total disease burden and the proportional attribution of total burden to specific causes and risk factors (*12*). By showing the disparities between diseases in the total burden, DALY estimates provide new insights that are potentially useful for the purposes of policy-making (*32*, *33*).

Conclusion

The main issue in burden of disease studies is access to complete, consistent and comparable epidemiological data. Summary measures of population health are only as good as the weakest link in the chain, which is the epidemiological evidence. The attempt to estimate burden of dementia in a cross-nationally comparable manner provided insight into the imperfections of the available data on prevalence, distribution of prevalence across severity classes, and excess mortality in dementia. Agenda-setting for the collection of epidemiological data is perhaps the most important issue to emerge from burden of disease estimation. ■

Acknowledgements

We thank Erik Nord (Institute of Public Health, Oslo, Norway) for help and advice in designing the valuation procedure, Theo Vos (Public Health Division, Department of Human Services, Melbourne, Australia) for help and advice on epidemiology and burden of disease estimation, Joshua A. Salomon (WHO, Geneva, Switzerland) for general advice, and Jan J. Barendregt (Department of Public Health, Erasmus Medical Center Rotterdam, Rotterdam, Netherlands) for help and advice with the epidemiological modelling.

The scientific work underlying this paper was performed at the centres that have participated in the European Disability Weights Project: National Institute of Public Health (formerly DICE), Copenhagen, Denmark; Health Services Management Centre and Department of Public Health and Epidemiology, University of Birmingham, Birmingham, England; Henri Mondor Hospital AP-HP, Paris, France; Department of Public Health, Erasmus Medical Center Rotterdam, Rotterdam, Netherlands; Institute of Social Medicine, Academic Medical Centre, University of Amsterdam, Netherlands; National School of Public Health, Carlos III Institute of Health, Madrid, Spain; and Department of Public Health Sciences, Karolinska Institute, Stockholm, Sweden.

The study was supported by a grant from the BIOMED II Programme of the European Union (project number BMH4-98-3253).

Conflicts of interest: none declared.

Résumé

Le projet European Disability Weights et la comparabilité inter-pays des estimations de la charge de morbidité

Objectif Rechercher l'origine de la variabilité des DALY (*disability-adjusted life years* = années de vie ajustées sur l'incapacité) observée entre différents pays dans le projet European Disability Weights.

Méthodes Le poids de l'incapacité correspondant à 15 états pathologiques a été calculé empiriquement dans cinq pays au moyen d'une méthode standardisée et les écarts entre pays des évaluations mesurées au moyen d'une échelle visuelle analogique ont été analysés. Comme exemple, on a estimé en DALY la charge morbide de la démence chez les femmes dans chacun des pays. On a analysé l'influence relative des variations d'un pays à l'autre de la démographie, de l'épidémiologie et du poids de l'incapacité sur les estimations des DALY.

Résultats La comparaison inter-pays des évaluations au moyen de l'échelle visuelle analogique a montré que l'ordre de classement est quasiment identique. Après standardisation des populations sur la taille et la structure d'âge, le taux de DALY pour 100 000 femmes s'établissait à 1050 en France et 1404 aux Pays-Bas. En raison de l'incertitude des données épidémiologiques, il est difficile d'estimer dans quelle mesure ces écarts représentent une variation vraie entre ces deux pays. L'utilisation des poids de l'incapacité européens plutôt que des poids spécifiques par pays n'a pas entraîné de modification importante des estimations de la charge morbide de la démence.

Conclusion L'estimation de la charge de morbidité et les comparaisons entre pays exigent avant tout des données épidémiologiques solides. Les estimations des DALY concernant la démence étaient relativement insensibles aux écarts du poids de l'incapacité entre pays européens.

Resumen

Comparabilidad interpaíses de las estimaciones de la carga de morbilidad: Proyecto Europeo sobre la Ponderación de las Discapacidades

Objetivo Investigar el origen de las diferencias interpaíses en los años de vida ajustados en función de la discapacidad (AVAD) en el Proyecto Europeo sobre la Ponderación de las Discapacidades.

Métodos Se calcularon empíricamente las ponderaciones de la discapacidad para 15 estados patológicos en cinco países por medio de un procedimiento normalizado, y se analizaron las diferencias interpaíses entre las puntuaciones arrojadas por la escala visual analógica (EVA). Para cada país, se calcularon los AVAD correspondientes a la carga de demencia entre las mujeres, utilizada como ejemplo ilustrativo. Se hizo un análisis de los efectos relativos de las diferencias interpaíses en cuanto a demografía, epidemiología y ponderación de las discapacidades sobre las estimaciones de los AVAD.

Resultados La comparación interpaíses de las puntuaciones obtenidas mediante la EVA reveló unas ordenaciones casi idénticas. Después de normalizar en función del tamaño y la estructura de edades de las poblaciones, las tasas de AVAD por 100 000 mujeres quedaron comprendidas entre los valores de 1050 para Francia y 1404 para los Países Bajos. Debido a la incertidumbre que afectaba a los datos epidemiológicos, es difícil determinar en qué medida esas diferencias reflejan una divergencia real entre los países. El empleo de ponderaciones de la discapacidad europeas en lugar de las propias de cada país no alteró sustancialmente las estimaciones de la carga de morbilidad impuesta por la demencia.

Conclusión La disponibilidad de datos epidemiológicos robustos es el primer requisito para poder estimar la carga de morbilidad y efectuar comparaciones pertinentes entre los países. Los AVAD estimados para la demencia fueron relativamente insensibles a las diferencias en las ponderaciones de la discapacidad entre los países europeos.

References

1. *The world health report 2000. Health systems: improving performance*. Geneva: World Health Organization; 2000.
2. Murray CJL, Lopez AD, editors. *The global burden of disease*. Cambridge (MA): Harvard School of Public Health on behalf of the World Health Organization and the World Bank; 1996. Global Burden of Disease and Injury Series, Vol. 1.
3. Melse JM, Essink-Bot ML, Kramers PGN, Hoeymans N. A national burden of disease calculation: Dutch disability-adjusted life-years. *American Journal of Public Health* 2000;90:1241-7.
4. Peterson S, Backlund I, Diderichsen F. [*Burden of disease in Sweden – a Swedish DALY calculation*.] Stockholm: National Public Health Institute; 1998. p. 50 (in Swedish).
5. Mathers C, Vos T, Stevenson C. *The burden of disease and injury in Australia*. Canberra: Australian Institute of Health and Welfare; 1999. Also available from: URL: www.aihw.gov.au.
6. Arnesen T, Nord E. The value of DALY life: problems with ethics and validity of disability adjusted life years. *BMJ* 1999;319:1423-5.
7. Ustün TB, Rehm J, Chatterji S, Saxena S, Trotter R, Room R, et al. Multiple-informant ranking of the disabling effects of different health conditions in 14 countries. *Lancet* 1999;354:111-5.
8. Jelsma J, Chivaura VG, Mhunwa K, De Weerdt W, De Cock P. The global burden of disease disability weights. *Lancet* 2000;335:2079-80.
9. Cooper RS, Osotimehin B, Kaufman JS, Forrester T. Disease burden in sub-Saharan Africa: what should we conclude in the absence of data? *Lancet* 1998;351:208-10.
10. Brooks RG. EuroQoL – the current state of play. *Health Policy* 1996;37:53-72.
11. Krabbe PFM, Stouthard MEA, Essink-Bot ML, Bonsel GJ. The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. *Journal of Clinical Epidemiology* 1999;52:293-301.
12. Murray CJL. Rethinking DALYs. In: Murray CJL, Lopez AD, editors. *The global burden of disease*. Cambridge (MA): Harvard School of Public Health on behalf of the World Health Organization and the World Bank; 1996. p.1-98.
13. Stouthard MEA, Essink-Bot ML, Bonsel GJ, and the Dutch Disability Weights Group. Disability weights for diseases – a modified protocol and results for a Western European region. *European Journal of Public Health* 2000;10:24-30.
14. Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-Economic Planning Sciences* 1976;10:129-36.
15. Streiner DL, Norman GW. *Health measurement scales: a practical guide to their development and use.* Oxford: Oxford University Press; 1995.

16. Lobo A, Launer LJ, Fratiglioni L, Andersen K, Di Carlo A, Breteler MM, et al. Prevalence of dementia and major subtypes in Europe: A collaborative study of population based cohorts. *Neurology* 2000;54 Suppl 5:S4-9.
17. Andersen K, Lolk A, Nielsen H, Andersen J, Olsen C, Kragh-Soerensen P. Prevalence of very mild to severe dementia in Denmark. *Acta Neurologica Scandinavica* 1997;96:82-7.
18. Saunders PA, Copeland JRM, Dewey ME, Gilmore C, Larkin BA, Phaterpekar H, et al. The prevalence of dementia, depression and neurosis in later life: the Liverpool MRC-ALPHA study. *International Journal of Epidemiology* 1993; 22:838-47.
19. Letenneur L, Gilleron V, Commenges D, Helmer C, Orgogonzo JM, Dartigues JF. *Journal of Neurology, Neurosurgery, and Psychiatry* 1999;66:177-83.
20. Witthaus E, Ott A, Barendregt J, Breteler M, Bonneux L. Burden of mortality and morbidity from dementia. *Alzheimer Disease and Associated Disorders* 1999;13:176-81.
21. Lobo A, Saz P, Marcos G, Dia J-L, De-la-Camara C. The prevalence of dementia and depression in the elderly community in a Southern European population. *Archives of General Psychiatry* 1995;52:497-506.
22. Fratiglioni L, Grut M, Forsell Y, Viitanen M, Winblad B. Clinical diagnosis of Alzheimer's disease and other dementias in a population survey. Agreement and causes of disagreement in applying DSM-III-R criteria. *Archives of Neurology* 1992;49:927-32.
23. Jorm AF, Korten AE, Jacomb A. Projected increases in the number of dementia cases for 29 developed countries: application of a new method for making projections. *Acta Psychologica Scandinavica* 1988;78:493-500.
24. Murray CJL, Lopez AD. Global and regional descriptive epidemiology of disability: incidence, prevalence, health expectancies and years lived with disability. In: Murray CJL, Lopez AD, editors. *The global burden of disease*. Cambridge (MA): Harvard School of Public Health on behalf of the World Health Organization and the World Bank; 1996. Global Burden of Disease and Injury Series, Volume 1. See also: URL: www.who.int/evidence/dismod.
25. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The Australian Burden of Disease Study: measuring the loss of health from diseases, injuries and risk factors. *Medical Journal of Australia* 2000;172;592-6.
26. Anand S, Hanson K. Disability-adjusted life years: a critical review. *Journal of Health Economics* 1997;16:685-702.
27. Williams A. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics* 1999;8:1-8.
28. Barendregt JJ, Bonneux L, Van der Maas PJ. DALYs: the age-weights on balance. *Bulletin of the World Health Organization* 1996;74:439-43.
29. Nunnally JC. *World health statistics annual*, 1995. Geneva: World Health Organization; 1996.
30. *Psychometric theory*. New York: McGraw-Hill; 1978.
31. Wilkins K, Parsons GF, Gentleman JF, Forbes WF. Deaths due to dementia: an analysis of multiple causes of death data. *Chronic Diseases in Canada* 1999;20:26-35.
32. Murray JL, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 1997;349:1436-42.
33. Gwatkin DR, Guillot M, Heuveline P. The burden of disease among the global poor. *Lancet* 1999;354:586-9.

Annex 1

Sources of data on prevalence of dementia

Denmark

Prevalence data were obtained from the Odense study in which 5237 eligible persons aged 65–84 years were randomly drawn from the Central Person Number Register (*17*). The sample included persons living in the community and persons living in institutions. Of this number, 3346 (64%) participated in the prevalence survey that was conducted between 1992 and 1994. The screening procedure included the Cambridge Examination for Mental Disorders in the Elderly Cognitive Subscale (CAMCOG) (including the Mini-mental State Examination). Persons giving a positive result were further examined with a standardized psychiatric interview, and 234 were found to have dementia.

England and Wales

Prevalence data were obtained from the MRC-ALPHA Study, which included subjects aged ⩾65 years living in the municipality of Liverpool (*18*); institutions were included. Samples were randomly selected from the general practitioner registry within strata of age (5-year bands) and sex of equal size. The baseline sample included 5222 subjects. The baseline survey was conducted in 1989 and 1990. A two-phase procedure was applied in order to determine whether a subject was demented or not. Phase 1 included the Geriatric Mental State, the Mini-mental State Examination and other tests. A sample of screen-positive subjects was subjected to a further interview by a psychiatrist using the Geriatric Mental State and CAMCOG.

France

The Paquid study was conducted in 75 parishes in the *départements* of Gironde and Dordogne (*19*). The sample was randomly selected from electoral rolls using a multistage procedure based on strata of age, sex, and the size of the geographical unit. The participants in the baseline prevalence survey were aged ⩾65 years and lived at home, i.e. persons in institutions were not included. The baseline sample included 3777 persons and the baseline examination started in 1988; 102 cases of dementia were found among 2792 subjects.

Netherlands

The Rotterdam study included all persons aged 55 years and over living in the district of Ommoord (*20*). The baseline sample comprised 7983 individuals surveyed in 1990, 494 of whom (1.2%) were identified as having dementia. The Mini-mental State Examination and Geriatric Mental Scale-A were used to screen the study population. Screen-positive persons underwent further testing on the basis of the Cambridge Examination for Mental Disorders in the Elderly (CAMDEX). The final step in case-finding involved neuroimaging and clinical examination by a neurologist and a neuropsychologist. Diagnosis was established on the basis of all available test results and other information by an expert panel using Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition (DSM-III-R) criteria. The classification of minimal, mild, moderate or severe dementia was based on clinical dementia ratings if available, otherwise on Mini-mental State Examination scores.

Spain

The Zaragoza study included 1134 individuals aged 65 years and over living in the urban area of Zaragoza, who were randomly selected with replacement in strata of age and sex from the municipal register (*21*). The baseline survey was conducted in 1988 and 1989. The screening test included the Mini-mental State Examination (cut-off point <24) and the Geriatric Mental State (cut off point >0).

Sweden

The Kungsholmen project included 2368 individuals aged ⩾75 years who were living in the Kungsholmen parish of Stockholm at baseline in October 1987 (including institutions) (*22*).

Annex 2

Jorm's model

Jorm's model allows for country-specific differences in dementia prevalence levels but assumes a similar course of dementia in different populations. Jorm observed an exponential increase in the prevalence of dementia with age across studies and proposed a formula based on all prevalence studies published up to 1987:

$$P_i = \exp (S_i - 13.50 + 0.137\, X),$$

where

P_i = prevalence rate of dementia for individuals aged X years from study i;

S_i = variable term, specific to study i, independent of age;

X = age in years.