Multiple-informant ranking of the disabling effects of different health conditions in 14 countries

*T Bedirhan Üstün, Jürgen Rehm, Somnath Chatterji, Shekhar Saxena, Robert Trotter, Robin Room, Jerome Bickenbach, and the WHO/NIH Joint Project CAR Study Group**

Summary

Background The Global Burden of Disease study provided international statistics on the burden of diseases, combining mortality and disability, that can be used for priority setting and policy making. However, there are concerns about the universality of the disability weights used. We undertook a study to investigate the stability of such weighting in different countries and informant groups.

Methods 241 key informants (health professionals, policy makers, people with disabilities, and their carers) from 14 countries were asked to rank 17 health conditions from most disabling to least disabling. Kruskal-Wallis ANOVA was used to test for differences in ranking between countries or informant groups and Kendall τ-B correlations to measure association between different rank orders.

Findings For 13 of 17 health conditions, there were significant ($p<0.05$) differences in ranking between countries; in the comparison of informant groups, there were significant differences for five of the 17 health conditions. The overall rank order in the present study was, however, almost identical to the ranking of the Global Burden of Disease study, which used a different method. Most of the rank correlations between countries were between 0·50 and 0·70 (average 0·61 [95% CI 0·59–0·64]). The average correlation of rank orders between different informant groups was 0·76.

Interpretation Rank order of disabling effects of health conditions is relatively stable across countries, informant groups, and methods. However, the differences are large enough to cast doubt on the assumption of universality of experts' judgments about disability weights. Further studies are needed because disability weights are central to the calculation of disability-adjusted life years.

Lancet 1999; **354:** 111–15
See Commentary page ???

*Members listed at end of paper

WHO, Assessment Classification and Epidemiology Group, Geneva (T Bedirhan Üstün MD, J Rehm PhD, S Chatterji MD, S Saxena MD, J Bickenbach PhD)**; Addiction Research Foundation, Centre for Addiction and Mental Health, Toronto, Canada** (J Rehm, R Room PhD)**; Department of Anthropology, Northern Arizona University, Flagstaff, AZ, USA** (R Trotter PhD)**; National Institute for Alcohol and Drug Research, Oslo, Norway** (R Room)**; and Queen's University, Kingston, Canada** (J Bickenbach)

Correspondence to: Dr T Bedirhan Üstün, Assessment Classification and Epidemiology Group, WHO, CH 1211 Geneva 27, Switzerland (e-mail: ustunb@who.ch)

Introduction

The Global Burden of Disease study[1] has attracted the attention of policy makers and public-health experts because the study provides a common unit for evaluation and priority setting for a wide range of health disorders. This unit, the disability-adjusted life year (DALY), adds disability to mortality in the overall estimation of the burden of disease. The addition of disability has increased the relative importance of non-communicable diseases, and the result is a more realistic measure of the global burden of diseases than that obtained from mortality alone.[2,3] DALYs provide a framework for different professions in determining priorities in health and human services and in evaluating the efficacy of interventions since they are based on life years, a universal measure.[3] Policy makers have a particular interest in making decisions on resource allocation and policy, and monitoring the impact of health-care reforms and other interventions with a common measure applicable in cost-effectiveness studies. The DALY is useful because it provides a common measure by making valid epidemiological data usable in an internally consistent approach.[4]

There have been various attempts to develop composite health measures that combine information on mortality and non-fatal health outcomes to represent population health in a single number.[5] One of the most widely adopted methods uses the quality-adjusted life years (QALY) model.[6,7] This model attempts to measure the impact of a disease in terms of different combinations of duration and quality of life.[8] DALY is the complement to QALY, the disability weight being the reverse scale of quality of life. Healthy life years (HLYs) and health-adjusted life years (HALYs) are other methods used to measure burden of disease.[9] The cross-cultural applicability of these methods and the equivalence of derived preferences have not been standardised globally, regionally, or nationally.

Although DALYs improve our understanding of international health statistics and can influence the policy process through decisions on resource allocation, great care must be taken in their construction, which involves choices and value judgments.[5] There is great concern that the calculation of the disability component in DALYs is based principally on disease-specific disability weights that indicate values for less than perfect health. That is, a perfectly healthy state has a weight of 0, and death is equivalent to a weight of 1. Hence, disease-related disability is placed between these endpoints: for example, angina has a weight of 0·223, major unipolar depression 0·619, and quadriplegia 0·895.[1] In the Global Burden of Disease study these weights were determined with professional health-care providers through the person trade-off method.[10] Professional health-care providers were chosen because they are familiar with health conditions and their outcomes, and such familiarity

makes it easier to form the often complex comparisons between the impacts of different disease states required by the person trade-off protocol.

Professional health-care providers were assumed to be representative of society's preferences with regard to resource allocations in health care. This assumption, however, requires empirical support, and for empirical testing, preference measures should also be obtained from various other groups such as policy makers and individuals with disabilities to see how these measures converge.

Furthermore, these disability weights are presumed to be universal, that is, equal across countries and cultures. Yet they were in fact derived from a single person trade-off exercise carried out at the WHO in Geneva. Although the results of other exercises since then suggest similar results,[1] there is a clear need for more systematic testing across different cultures and different informant groups with other forms of measurement. This study was motivated by these concerns about the universality of the disability weights used in the construction of DALYs.

In a large study on the cross-cultural applicability of a proposed revision of the International Classification of Impairments, Disabilities and Handicaps,[11,12] a sub-study examined whether expert ratings on the disability effects of different health conditions were universal, in the sense of being stable across methods, cultures, and informant groups. Ranking was the chosen method rather than the person trade-off method because ranking requires less specialised participants, less time, and no technical knowledge. The original Global Burden of Disease study protocol also used an ordinal ranking exercise in addition to a variant of the person trade-off protocol, and required respondents to reconcile the discrepancies between the two during a deliberative phase.[1] Our study attempted to replicate independently the results of the exercise carried out within the framework of the Global Burden of Disease study in different cultures with different informants. We should point out, however, that because of the difference in methods, our study can only test the assumptions of the Global Burden of Disease study (eg, stability of disablement scores across cultures and informant groups). With the ranking exercise, disability weights cannot be derived. The choice of method is justified since the main aim of our study is to test the underlying assumptions of the Global Burden of Disease study. Ranking exercise also requires much less time (10–15 min instead of 2 days per group for the person trade-off method).

We aimed to answer four questions. Are there significant differences in the ranking of the disabling health conditions by key informants from different countries? Are there significant differences in the ranking of the disabling health conditions by respondents from different informant groups (medical professionals, allied health professionals, health-policy makers, consumers, or caregivers)? Could the ranking of the disability weights of the Global Burden of Disease study be replicated with a different method? What are the underlying patterns in respondent ratings of disabling effects of health conditions?

Methods

Participants

Informants from 14 countries took part in this study (Canada, China, Egypt, Greece, India, Japan, Luxembourg, the Netherlands, Nigeria, Romania, Spain, Tunisia, Turkey, and the UK). Thus, all WHO regions were represented in the study.

Key informants were defined as those who, by virtue of their position and knowledge, have an understanding of disabilities that makes them representative spokespersons for their culture. For each country, 15 informants were selected—three individuals from each of five groups: medical professionals (eg, physician, psychiatrist, psychologist, nurse); allied health professionals (eg, social worker, case worker); policy makers or opinion leaders in the area of disability services; individuals with a disabling physical health condition, and their caregivers; and individuals with a disabling health condition related to alcohol, drugs, or mental health, and their caregivers. The final number of participants included in this study varied slightly by country, but essentially, the quota of 15 was achieved in each country.

Design

Informants responded to a questionnaire including both open-ended items and closed-ended items. In rank-ordering disablements, key informants were presented with a deck of 17 cards listing different health conditions with a short description. They were instructed to rank conditions from the most disabling to the least disabling. The most disabling condition was described as one that would make activities such as dressing, feeding, moving around, and meeting basic day-to-day responsibilities very difficult. The least disabling conditon was described as one that would not interfere with the above-mentioned activities in day-to-day life. Codes were assigned to the 17 conditions, with 01 representing the most disabling, and 17 representing the least disabling. 13 of the 17 health conditions were adapted from the 22 indicator conditions of the Global Burden of Disease study to measure the burden of disabilities and diseases within the DALY framework. The selection of the original indicator conditions was based on results from pretests, which showed that respondents in certain cultures found some diseases difficult to grasp. Ranking of 22 conditions has also been shown to be too much of a burden on the respondent.[13] In addition to the original conditions, we included disorders related to alcohol misuse and drug misuse, and HIV infection, which we assumed would rank as medium disabling health conditions and would show cultural differences.[14]

The ranking exercise was carried out through an interview that lasted, on average, 90 min. The interviewer showed the informant a collection of demographic data, and asked questions about the language of disablement in the culture, about existing compensation systems for disablement, about the social stigma of disabilities, and about other societal reactions to disabling health conditions. This process served to frame the issue of the ranking exercise within the context of understanding the disabling consequences of health conditions. The assessment of social disapproval or stigma for different health and social conditions was quantified on an 11-point rating scale with the endpoints labelled as none (0) and extreme (10).

Each key informant was presented with a brief overview of the interview and its purpose. The interviewer then asked questions about demographic variables of the informant and about their experience in the area of disabilities. The interviewer also asked open-ended questions on language and disablement and identification of and societal reactions to disablement associated with selected conditions (eg, difficulties with walking, psychotic symptoms, low intelligence, and alcohol-related and drug-related problems). Once the open-ended portion of the interview was completed, the interviewer presented the key informant with the 17 different cards, and administered the instructions to rank the conditions in order according to their disabling effects. Finally, the key informant was asked to complete a self-administered questionnaire that addressed expected difficulties with day-to-day activities in a selection of health conditions and included the assessment of stigma related to various disabilities.

Health condition	Rank	Median	Mean (SD)	n
Quadriplegia	1	2	3·3 (3·2)	241
Dementia	2	4	4·9 (3·6)	241
Active psychosis	3	4	5·3 (3·6)	241
Paraplegia	4	5	5·9 (3·3)	241
Blindness	5	6	6·8 (4·0)	241
Major depression	6	6	7·2 (3·8)	241
Drug dependence	7	8	7·8 (3·9)	222
HIV infection	8	9	8·8 (5·2)	239
Alcoholism	9	9	9·2 (3·6)	241
Total deafness	10	10	9·4 (3·7)	241
Mild mental retardation	11	10	9·9 (3·6)	241
Incontinence	12	10	10·2 (4·1)	239
Amputation below the knee	13	11	10·2 (3·7)	241
Rheumatoid arthritis	14	12	11·5 (3·6)	241
Severe migraines	15	12	11·6 (3·8)	240
Infertility	16	16	14·6 (3·6)	238
Vitiligo on face	17	16	15·0 (2·4)	238

n=number of informants.

Table 1: **Rank order of disabling effect of health conditions by severity**

Statistical analyses

Non-parametric statistics for ordinal variables were used to analyse the data. Overall ranking was established on the basis of median of the ranks. For conditions with the same median, the arithmetic mean of rankings was taken as a second criterion.

To test for differences between countries or between informant groups, the Kruskal-Wallis rank-order ANOVA for one factor was used. Kendall τ-B correlations were used to measure the association between different rank orders. This measure was chosen instead of Spearman's rank correlation coefficient because it has a clear interpretation of choice: Kendall τ is equivalent to the number of times rankings agree about a pair, minus the number of times rankings disagree, divided by the total number of pairs. Thus, τ can be interpreted as the corresponding difference between probabilities of agreement and disagreement with respect to the ranking of each pair.[15] Statistical analyses were calculated with SYSTAT 8.0 and with StatXact 3.1 (providing exact calculations of the Kendall τ correlation and the associated CIs).

Results

Interview data were collated, and the 17 health conditions were ranked from most disabling to least disabling (table 1). Overall, the ranking of the conditions at both ends of the range showed less variability between informants than that for intermediate conditions.

There were, however, deviations from this order across countries (table 2). Active psychosis was ranked most disabling in the Netherlands and Canada but third overall. HIV infection was given a low ranking by informants from Japan, Luxembourg, Spain, Turkey, and the UK, whereas it was ranked first in Egypt and Tunisia. HIV infection was the health condition with the most variation in ranking.

The differences between countries were significant ($p<0·05$) for 13 of the 17 health conditions on the Kruskal-Wallis test. Only quadriplegia, paraplegia, amputation below the knee, and mild mental retardation did not show differences in ranking between countries at $p<0·05$. Three of the four conditions that were judged uniformly across countries were physical disabilities. The fourth, mild mental retardation, also showed little variation though this condition was less uniformly rated than the other three.

Although there were significant differences in the ranking between countries, there is evidence of convergence in judgments. The average of the Kendall τ rank correlations between different countries was 0·61 (95% CI 0·59–0·64). Within this average, there were, however, clear cultural differences for some comparisons: for example, τ for the Netherlands and Turkey was 0·41 (0·14–0·69), whereas that for the Netherlands and Tunisia was 0·29 (0·03–0·56). For Luxembourg and Spain τ was 0·87 (0·74–0·99), and for Romania and India 0·82 (0·72–0·92). Most of rank correlations ranged between 0·50 and 0·70 (for $\tau=0·50$, typical 95% CI 0·27–0·73; for $\tau=0·60$, typical 95% CI 0·35–0·85). A table with all bivariate correlations between countries can be obtained from the corresponding author. Overall, these correlations did not support a general notion of universal disability weights.

Only five of 17 health conditions had significantly different rank orders between different informant groups: quadriplegia, HIV infection, total deafness, mild mental retardation, and amputation below the knee. Again, most of these were physical conditions, but in this analysis they showed most variation, whereas in the analysis by country they showed least variation.

The rank orders showed significant differences between some informant groups that might have been expected to have convergent views (table 3). For example, the ratings between health professionals working in physical rehabilitation showed the largest differences from consumers and caregivers in the same sector with respect to rankings in the above

Overall rank*	Rank order by country													
	Canada (n=15)	China (n=16)	Egypt (n=15)	Greece (n=15)	India (n=43)	Japan (n=18)	Luxembourg (n=16)	Netherlands (n=13)	Nigeria (n=15)	Romania (n=15)	Spain (n=18)	Tunisia (n=15)	Turkey (n=15)	UK (n=12)
1	2	1	2	1	1	2	1	3	1	1	1	2	1	2
2†	3	8	3	3	2	1	2	2	6	2	2	3	2	1
3†	1	5	4	2	5	3	3	1	3	3	4	6	4	4
4	4	4	8	4	4	5	7	7	2	4	5	4	3	5
5†	8	3	5	9	3	4	4	9	5	5	6	5	5	8
6†	5	6	7	7	6	8	6	4	4	7	3	7	11	3
7†	7	2	6	6	11	7	5	6	10	11	8	11	7	NA
8†	10	9	1	5	7	13	15	12	8	8	13	1	14	14
9†	9	10	11	8	10	10	8	5	13	13	7	12	10	6
10†	11	12	10	13	9	6	9	11	15	9	9	13	12	12
11	6	11	9	12	12	15	10	13	11	10	10	9	8	7
12†	15	13	13	10	8	14	13	15	7	6	12	10	6	11
13	12	7	12	11	14	9	11	14	12	12	11	8	9	13
14†	14	14	17	15	13	11	14	10	14	15	15	16	13	10
15†	13	15	16	14	15	12	12	8	9	14	14	17	15	9
16†	16	17	14	16	17	16	17	16	16	16	17	15	17	16
17†	17	16	15	17	16	17	16	17	17	17	16	14	16	15

NA=data not collected. *See table 1. †Significant differences between countries at $p<0·05$ (Kruskal-Wallis rank-order ANOVA).

Table 2: **Disability ranks associated with different health conditions by country**

Health condition	Overal rank	Medical professionals treating:					Allied health professionals (n=51)	Health-policy makers (n=35)	Consumers and carers	
		Physical illness (n=14)	ADM disorders (n=35)	Physical and ADM disorders (n=14)	Other (n=11)	All (n=74)			Physical disability (n=30)	Mental disability (n=45)
Quadriplegia	1	1	1	3	1	1	1	1	1	1
Dementia	2	2	2	1	2	2	3	2	4	2
Active psychosis	3	5	3	2	4	3	2	3	2	3
Paraplegia	4	3	4	4	3	4	5	4	6	4
Blindness	5	6	5	5	5	5	6	7	5	5
Major depression	6	4	6	6	6	6	4	5	8	6
Drug dependence	7	9	7	7	8	7	7	9	7	8
HIV infection	8	10	11	15	11	11	12	6	3	7
Alcoholism	9	13	10	10	9	9	8	10	9	9
Total deafness	10	12	8	9	7	8	10	11	11	12
Mild mental retardation	11	14	13	8	12	13	9	8	10	13
Incontinence	12	7	12	14	13	12	11	12	12	11
Amputation below the knee	13	8	9	12	10	10	13	13	13	10
Rheumatoid arthritis	14	11	14	11	15	14	14	14	15	14
Severe migraines	15	15	15	13	14	15	15	15	14	15
Infertility	16	16	16	17	16	16	16	17	17	16
Vitiligo on face	17	17	17	16	17	17	17	16	16	17

ADM=Alcohol, drug, and mental disorders.

Table 3: **Disablement ranks of different health conditions by informant group**

conditions, and physical conditions in general. Overall, the rank orders for different informant groups had an average correlation of 0·76.

The research design allowed us to measure the influence of social disapproval or stigma on disability rankings for a subset of the conditions. Higher degrees of stigma were associated with higher disability rankings; however, the effect size was small (<0·10 on average, for τ; data not shown). We conclude that stigma, as measured on an 11-point rating scale, has a measurable but very small effect on disability rankings.

The resulting summary ranking across all judgments was very similar to the ranking by experts in the Global Burden of Disease study (table 4), derived by a different method.[17] The rank order correlation (τ) was 0·77 (0·40–1·00; Z=3·66, $p<0·01$). The only notable difference was for severe migraine, which ranked third in the Global Burden of Disease study, and 11th in this study. However, the way judgments were elicited for migraine in the two exercises differed. In the Global Burden of Disease study, experts were repeatedly reminded that they should consider the artificial case of 1 continuous year of severe migraines, with the consequence of staying in bed most of the time, when making their judgments. We did not repeatedly remind our informants about this artificial case; thus, the comparability of the condition severe migraine may have been compromised. With severe migraine omitted, τ was 0·97 (0·89–1·00; Z=4·39; $p<0·01$).

Health condition	Rank order GBD study*	Rank order from this study	Average disability weights GBD	Median disability weights GBD
Quadriplegia	1	1	0·895	0·894
Dementia	2	2	0·762	0·770
Active psychosis	4	3	0·722	0·714
Paraplegia	5	4	0·671	0·672
Blindness	6	5	0·624	0·642
Major depression	7	6	0·619	0·633
Total deafness	9	7	0·333	0·333
Mild mental retardation	8	8	0·361	0·410
Below the knee amputation	10	9	0·281	0·272
Rheumatoid arthritis	11	10	0·209	0·200
Severe migraines	3	11	0·738	0·753
Infertility	12	12	0·191	0·177
Vitiligo on face	13	13	0·020	0·005

*Taken from table 1·3 (page 39) of the Global Burden of Disease Study (GBD).[1]

Table 4: **A comparison of the rank order between Global Burden of Disease study[1] and this study**

Discussion

Although there was some variation, the rankings of the disabling effect of health conditions are stable across countries and informant groups, irrespective of the method used for the ranking, as shown by the high correlations between the Global Burden of Disease study and this study, and the close agreement among the 14 countries and among the eight informant groups in our study. Thus, the responses from the key informants show that the relative burden of different health conditions, in terms of disability, is fairly similar across the world. The results also, however, indicate that there are systematic and, in some cases, pronounced differences between cultures and informant groups. These differences are large enough to shed doubt on the assumption of universality of the disability rankings, and subsequently on the weights, and they are large enough to be investigated further in a systematic way.

In theory, there is no reason why disability weights should be universal. First, the actual burden of disability or activity limitations is affected by many factors; for example, the burden of disability for individuals with paraplegia is likely to depend on the assistance available in the form of devices (wheelchairs, adapted cars, adapted workplaces) and social support. There are clear differences between countries in the availability of assistance and these differences should be reflected in the disability weights attached to certain health conditions. Second, valuation and description in the terminology of the Global Burden of Disease study may not be strongly related. Ranking exercises for this study were directed exclusively towards the extent of disabling conditions. All diseases had to be ranked solely on the basis of their disabling effects. However, the person trade-off method asks for a valuation (ie, trade-off of people with that disease). The basic comparisons are obtained through questions about the relative weights attached to the lives of 1000 healthy people for 1 year or the deaths of 2000 blind people. Thus, issues other than the disabling effects of health conditions, such as prognosis, pain, mood impact, and public opinon may drive the valuations.

The condition ranked most variably by country in our sample was HIV infection; the differences between expert groups are less dramatic but still important. Awareness and availability of information on and treatment for HIV infection in different countries vary greatly; also the infection carries a stigma since it is generally transmitted sexually. Physical conditions were ranked more uniformly than were mental conditions across countries, but not across informant groups. This finding implies that, though physical conditions such as quadriplegia, paraplegia, and amputation are, as a group, viewed differently from conditions such as active psychosis, these differences are more likely to be a feature of the individual respondent's characteristics than cultural differences. Furthermore, the conditions with the most variation are mental conditions that lie in the mid-range of the rankings.

Any ordinal ranking can only indicate the relative effects of disability from different health conditions (eg, active psychosis causes more disability than major depression), and not the absolute effects, which may vary greatly between countries because of different formal and informal health and social-support systems. Also, close agreement between health professionals or other key informants does not necessarily indicate that the real disability associated with the different health conditions is similarly uniform across countries. In different instances in the past, the consensus among experts has been at variance with the actual behaviour of patients.[16,17] Thus, the next step should be empirical studies to examine whether selected health conditions really do have the same disabling effects in different countries and cultures across the world, and whether they are stable with respect to different assessment methods and different informant groups.

Finally, an actual measurement of functioning of patients with different health conditions in their respective contexts, ought to be coupled with valuation exercises across cultures and respondent groups.[18] This measurement may then provide valuable insights into the determinants of values that people assign to disabling health conditions. If public policy decisions, such as resource allocation and priority setting, are to be based on evidence, we need to make sure that they are formed by all parties and are based on real-life data obtained in accordance with high scientific standards.

Contributors

T Bedirhan Üstün was the principal investigator and contributed to all parts of the study. J Rehm codesigned the study, was responsible for the statistical analyses, and contributed to the writing of the paper. S Chatterji codesigned the study, and contributed to the interpretation of results. S Saxena codesigned the study, did part of the fieldwork, and contributed to the writing of the paper. R Trotter codesigned the study, conceptualised the qualitative parts, and analysed them. R Room took part in the design and the analysis of the study and participated in the revision. J Bickenbach took part in the interpretation and writing of the results, and in the revision of the article.

The WHO/NIH Joint Project CAR study advisers and principal investigators

R Battjes (NIDA, USA), B Grant (NIAAA, USA), C Kennedy (NIMH, USA), D Regier (NIMH, USA), S Yu Cun (China), V Mavreas (Greece), R S Murthy (Bangalore, India), H Pal (Delhi, India), R Thara (Chennai, India), M Tazaki (Japan), C Pull (Luxembourg), H Hoek (Netherlands), A Odejide (Nigeria), R Vrasti (Romania), J L Vazquez-Barquero (Spain), A Chaker (Tunisia), A Gogus (Ankara, Turkey), N Dedeoglu (Antalya, Turkey), K Ogel (Istanbul, Turkey), D Mumford (UK).

Acknowledgments

This study was made possible within the WHO-NIH Joint Project (UO1MH35883).

References

1 Murray CJL, Lopez AD. The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Cambridge, MA: Harvard University Press, 1996.

2 Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020. Global Burden of Disease Study. *Lancet* 1997; **349:** 1498–504.

3 Murray CJL, Lopez AD. Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: Global Burden of Disease Study. *Lancet* 1997; **349:** 1347–352.

4 Ad Hoc Committee on Health Research Relating to Future Intervention Options: investing in health research and development. Geneva: WHO, 1996.

5 Field MJ, Gold MR. Summarising population health: directions for the development and application of population metrics. Institute of Medicine, Washington DC: National Academy Press, 1998.

6 Torrence GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986; **5:** 1–30.

7 Froberg DG, Kane RL. Methodology for measuring health-state preferences-I: measurement strategies. *J Clin Epidemiol* 1989; **42:** 345–54.

8 Mehraz A, Gafni A. Preference-based outcome measures for economic evaluation of drug interventions: quality adjusted life years (QALYs) versus healthy years equivalents (HYEs). *Pharmacoeconomics* 1992; **1:** 338–45.

9 Hyder AA, Rottlant G, Morrow RH. Measuring the burden of disease: healthy life years. *Am J Public Health* 1998; **88:** 196–202.

10 Nord E. The person-trade off approach to valuing health care programs. *Med Decis Making* 1995; **15:** 201–08.

11 WHO. International classification of impairments, disabilities and handicaps. Geneva: WHO, 1980.

12 WHO. International classification of impairments, activities and participation (ICIDH-2) Beta-1 draft. Geneva: WHO, 1997.

13 Trotter RT, Üstün TB, Chatterji S, et al. Cross-cultural applicability research on disablement: models, methods and contributions to revision of the international classification. *Hum Org* (in press).

14 Mäkelä K, Room R, Single E, et al. Alcohol, society and the state-I, a comparative study of alcohol control. Toronto: Addiction Research Foundation, 1981.

15 Hays WL. Statistics for the social sciences. New York: Holt, Rinehart & Winston, 1973.

16 Rehm J, Gadenne V. Intuitive predictions and professional forecasts: cognitive processes and social consequences. Oxford: Pergamon Press, 1990.

17 Single E. The concept of harm reduction and its application to alcohol: the sixth annual Dorothy Black lecture. *Drugs: Ed, Prev Pol* 1997: **4:** 7–22.

18 Üstün TB, Chatterji S. Editorial: measuring functioning and disability—a common framework. *Int J Meth Psychiatr Res* 1998; **7:** 79–83.