Disability weights for the Global Burden of Disease 2013 study

Disability weights for the Global Burden of Disease 2013 study

Joshua A Salomon, Juanita A Haagsma, Adrian Davis, Charline Maertens de Noordhout, Suzanne Polinder, Arie H Havelaar, Alessandro Cassini, Brecht Devleesschauwer, Mirjam Kretzschmar, Niko Speybroeck, Christopher J L Murray, Theo Vos

Summary

Background The Global Burden of Disease (GBD) study assesses health losses from diseases, injuries, and risk factors using disability-adjusted life-years, which need a set of disability weights to quantify health levels associated with non-fatal outcomes. The objective of this study was to estimate disability weights for the GBD 2013 study.

Methods We analysed data from new web-based surveys of participants aged 18–65 years, completed in four European countries (Hungary, Italy, the Netherlands, and Sweden) between Sept 23, 2013, and Nov 11, 2013, combined with data previously collected in the GBD 2010 disability weights measurement study. Surveys used paired comparison questions for which respondents considered two hypothetical individuals with different health states and specified which person they deemed healthier than the other. These surveys covered 183 health states pertinent to GBD 2013; of these states, 30 were presented with descriptions revised from previous versions and 18 were new to GBD 2013. We analysed paired comparison data using probit regression analysis and rescaled results to disability weight units between 0 (no loss of health) and 1 (loss equivalent to death). We compared results with previous estimates, and an additional analysis examined sensitivity of paired comparison responses to duration of hypothetical health states.

Findings The total analysis sample consisted of 30 230 respondents from the GBD 2010 surveys and 30 660 from the new European surveys. For health states common to GBD 2010 and GBD 2013, results were highly correlated overall (Pearson's r 0·992 [95% uncertainty interval 0·989–0·994]). For health state descriptions that were revised for this study, resulting disability weights were substantially different for a subset of these weights, including those related to hearing loss (eg, complete hearing loss: GBD 2010 0·033 [0·020–0·052]; GBD 2013 0·215 [0·144–0·307]) and treated spinal cord lesions (below the neck: GBD 2010 0·047 [0·028–0·072]; GBD 2013 0·296 [0·198–0·414]; neck level: GBD 2010 0·369 [0·243–0·513]; GBD 2013 0·589 [0·415–0·748]). Survey responses to paired comparison questions were insensitive to whether the comparisons were framed in terms of temporary or chronic outcomes (Pearson's r 0·981 [0·973–0·987]).

Interpretation This study substantially expands the empirical basis for assessment of non-fatal outcomes in the GBD study. Findings from this study substantiate the notion that disability weights are sensitive to particular details in descriptions of health states, but robust to duration of outcomes.

Funding European Centre for Disease Prevention and Control, Bill and Melinda Gates Foundation.

Introduction

The Global Burden of Disease (GBD) study is a systematic scientific effort to compare the magnitude of health losses associated with different diseases, injuries, and risk factors worldwide. Health losses are quantified in the GBD study with disability-adjusted life-years, a summary measure of population health that combines information about premature mortality, expressed as years of life lost, with information about non-fatal health losses attributable to time lived in states worse than full health, expressed as years lived with disability. To compute years lived with disability for a particular health outcome in a given population, the number of people living with that outcome is multiplied by a disability weight that represents the magnitude of health loss associated with the outcome. Disability weights are measured on a scale from 0 to 1, with 0 implying a state that is equivalent to full health and 1 a state equivalent to death.

The conceptual and methodological basis for estimation of disability weights in the GBD study has changed with various iterations of the study,[1–4] and some key issues relating to disability weights have been widely debated in the scientific literature.[5–19] For the GBD 2010 study,[20–21] a new approach to computation of disability weights was developed, with an emphasis on eliciting of weights that were consistent with the GBD measurement construct of health loss as opposed to the broad construct of welfare loss. To this end, in the GBD 2010 disability weights measurement study,[21] we developed a new set of standardised methods for primary data collection from general public survey respondents in a diverse array of settings, and an associated set of data analytic approaches to enable

Lancet Glob Health 2015; 3: e712–23

See **Comment** page e661

Department of Global Health and Population, Harvard T H Chan School of Public Health, Boston, MA, USA (Prof J A Salomon PhD); **Department of Public Health, Erasmus Medical Center, Rotterdam, Netherlands** (J A Haagsma PhD, S Polinder PhD); **Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA** (J A Haagsma, Prof C J L Murray MD, Prof T Vos PhD); **Public Health England, London, UK** (Prof A Davis PhD); **Institute of Health and Society, Université catholique de Louvain, Belgium** (C Maertens de Noordhout MD, B Devleesschauwer PhD, Prof N Speybroeck PhD); **Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA** (A H Havelaar PhD); **Institute for Risk Assessment Sciences, Utrecht University, Utrecht, Netherlands** (A H Havelaar); **European Centre for Disease Prevention and Control, Stockholm, Sweden** (A Cassini MPH); **Department of Virology, Parasitology and Immunology, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium** (B Devleesschauwer); **Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands** (Prof M Kretzschmar PhD); **and National Institute for Public Health and the Environment, Centre for Infectious Disease Control, Bilthoven, Netherlands** (Prof M Kretzschmar)

Correspondence to:
Prof Joshua A Salomon, Department of Global Health and Population, Harvard T H Chan School of Public Health, Boston, MA 02115, USA
jsalomon@hsph.harvard.edu

For the **Global Burden of Disease** see http://www.healthdata.org/gbd

Research in context

Evidence before this study

Before this study, the Global Burden of Disease (GBD) 2010 disability weights measurement study constituted the largest empirical effort to quantify health losses associated with a wide range of non-fatal effects of disease and injury causes. The 2010 disability weights measurement study introduced a new measurement approach relying on simple questions asked to respondents in the general community about which of two hypothetical people, each characterised by functional health levels and symptoms typical of a particular health condition, was regarded as being healthier than the other was. Key methodological issues raised by that study were the design of the lay descriptions used to describe each pair of health outcomes to respondents and the extent to which the resulting weights were sensitive to the choice of respondent samples.

Added value of this study

This study adds to the existing scientific literature for disability weights by doubling the empirical database for estimation of disability weights for the GBD study and exploring several key methodological issues at the frontier of the measurement agenda for disability weights. Specifically, this study aimed to examine the robustness of disability weights estimates to variation in a range of different elements, including aspects of the descriptions that are used to present outcomes to survey respondents; the composition of survey respondents from different settings; and the framing of disability weights questions in terms of either temporary or chronic outcomes.

Implications of all the available evidence

The new disability weights presented in this study provide quantitative estimates of health losses associated with non-fatal outcomes in GBD 2013. These weights answer questions about the comparative magnitude of health reductions due to a broad array of health outcomes and also supply an essential resource for clinicians and researchers investigating disease burden and policy makers who apply the results from these investigations. Findings from this study substantiate the plausibility of one important assumption in disability-adjusted life-years, which is that the comparative magnitude of health losses associated with different states might be stable over different durations of outcomes.

translation of these survey data into disability weights for all non-fatal outcomes included in the GBD.[21]

Data were collected in the GBD 2010 disability weights measurement study through a household survey in five countries and an open access web-based survey. The primary instrument for data collection was a simple paired comparison question in which respondents were asked to consider two hypothetical individuals, each with a particular health condition described briefly in terms of its main functional effects and symptoms, and to state which of the two individuals they would deem as being healthier than the other. An additional type of question, included on the web-based survey only, asked respondents to compare the population health benefits in two different hypothetical health programmes, and this information was used to anchor the results from the paired comparison data such that all weights were located on a scale between 0 and 1. The study produced a set of new disability weights, including measures of uncertainty for the 220 unique health states that collectively captured variation across the full set of 1160 sequelae arising from the 289 disabling disease and injury causes in GBD 2010. These disability weights were incorporated into estimates of years lived with disability, disability-adjusted life-years, and a related summary measure of average population health called healthy life expectancy.[22–24]

An important limitation acknowledged in the GBD 2010 disability weights measurement study was the crucial dependence of the results on the ways that outcomes were described to survey respondents.[21] Presentation of health states in terms of brief lay descriptions necessitated a focus on the most salient and typical effects on someone living in a particular health state. With the aim to balance validity and parsimony, descriptions invariably omitted some aspects of health states. Moreover, lay descriptions were developed collaboratively with individual expert groups—each organised around a particular set of health issues—and although efforts were made to ensure consistency across the set of conditions, some variability in language and detail remained. Commentators on the GBD 2010 disability weights measurement study have queried some of the results and suggested some specific examples of aspects of the lay descriptions that could have introduced imbalance across different conditions or inconsistency across severity levels for a given condition.[25–27]

In GBD 2013,[28] an expanded list of 306 disease and injury causes gives rise to 2337 sequelae, which have been mapped to 235 unique health states. In this study, we describe estimation of an updated set of disability weights for GBD 2013 via new data collection in four European countries and a reanalysis of existing and new data combined. In addition to updating the weights used for health states in GBD 2010 on the basis of an expanded set of survey data, the present study had three further objectives. The first objective was to estimate disability weights for additional sequelae that have been incorporated into GBD 2013 but were not included in GBD 2010. The second objective was to estimate revised disability weights for specific sequelae for which GBD 2010 lay descriptions lacked mention of an important symptom or consistency of wording across different levels of severity for the same impairment. The final objective was to assess one of the key assumptions in the

GBD disability weight measurement approach, which is that comparative assessments of health states are insensitive to the duration of these outcomes.

Methods

Study design and participants

The dataset for analysis in this study combined data collected in the previous surveys[21] done for the GBD 2010 disability weights measurement study and a set of four new surveys done as part of the European disability weights measurement study.[29] In the first component of GBD 2010, face-to-face household surveys were done in Bangladesh, Indonesia, Peru, and Tanzania in 2009–10 in samples designed to be representative of a specific geographical area in each country, and a computer-assisted telephone survey was done in the USA, designed to be nationally representative. Sampling in the surveys used a multistage, stratified design, with probabilities of being selected proportional to population size. In the second component of GBD 2010, we ran a web-based survey, with recruitment of respondents through various channels, including news items and commentaries in journals, announcements at scientific meetings, postings on relevant websites, online media campaigns, and direct outreach to members of the professional networks of the study investigators. In both the household and web-based surveys, eligible participants were aged 18 years or older.

For the present study, we collected new data in surveys that ran between Sept 23, 2013, and Nov 11, 2013, in Hungary, Italy, the Netherlands, and Sweden. The surveys were initiated in connection with a project sponsored by the European Centre for Disease Prevention and Control (the Burden of Communicable Diseases in Europe project),[30] and the four countries were selected as representing four regions of Europe (east, south, middle, and north) with respect to age, sex, and education. Samples were drawn from existing internet panels in each country, which are standing panels of respondents who take part in surveys administered via the internet. We recruited respondents from each panel on the basis of quota sampling with reference to age, sex, and education (categorised by level of schooling completed) to maintain population representativeness with respect to these individual characteristics. In the Netherlands, we pre-selected respondents before inviting them to participate, which was possible because age, sex, and educational level were already known for the panel members; in the other three countries, we invited panelists to participate via a website link and then selected them on the basis of their individual characteristics. All eligible participants were aged 18–65 years.

Procedures

We designed the European disability weights measurement study[29] to follow the protocol that was developed and applied in the GBD 2010 disability weights measurement study.[21] Specifically, the surveys relied mainly on a paired comparison task that asked respondents to consider descriptions of two hypothetical people, each with a particular health state, and specify which person they regarded as being healthier than the other. We presented health states with brief lay descriptions that focused on the major functional effects and symptoms associated with each health state (the appendix has a complete listing of all lay descriptions used in this study). As a secondary measurement task, some versions of the survey included a population health equivalence question, which asked respondents to compare the health benefits of different life-saving or health-improving programmes.

See Online for appendix

The European disability weights measurement study covered 255 health states, of which 183 have been used in GBD 2013. These 183 consisted of 135 of the 220 health states in GBD 2010 that were included in the European disability weights measurement study with unmodified lay descriptions; 30 from GBD 2010 for which alternative lay descriptions were included in the European disability weights measurement study; and 18 new ones that were added to GBD 2013 but not included in GBD 2010. The 18 new states include five mild health states for alcohol and drug dependence outcomes; two for epilepsy, aligning with the epidemiological data that define severe epilepsy in people who have, on average, one or more fits per month, and less severe epilepsy in those with between one and 11 fits in the past year; two mild health states for low back pain; two for treated or untreated amputation of one arm; and one state each for stress incontinence, concussion, hypothyroidism, hyperthyroidism, thrombocytopenic purpura, vertigo, and borderline intellectual functioning. We developed the lay descriptions for these new states using the same design principles used in GBD 2010, with a focus on description of the most salient attributes of a health state parsimoniously using simple, non-clinical vocabulary.

For the 30 GBD 2010 health states that were included with modified lay descriptions in the European disability weights measurement study, with resulting weights used in GBD 2013, we developed the modified descriptions in response to identified issues in consistency of presentation across different related states (for four levels of intellectual disability and seven states relating to amputation) or identified instances in which the original descriptions omitted key components of the functional characterisation of the outcome. These omissions led to development of alternative descriptions that added reference to incontinence for four spinal cord lesion health states; added psychological effects of social isolation for the three most severe hearing loss health states, and another three for those severities with inclusion of ringing in the ears, with revisions to the remaining hearing loss states to preserve consistency in presentation and ordering of these states; and incorporated minor revisions of the lay descriptions for mild depression, fracture of face bone, the three severity

levels for motor and cognitive impairment, and the generic health state for the worry associated with having a diagnosis of a chronic disease even if no symptoms are present. For these 30 states, the European disability weights measurement study included both the original and modified lay descriptions in the set of outcomes that were randomly allocated to respondents in pairs, with the restriction that no pairs would consist of both the original and modified description for the same state.

We randomly assigned (using a computer-generated random number sequence) respondents to one of three different versions of the web-based survey. One version (assigned to about 40% of respondents) included 15 paired comparison questions (drawn randomly from all available possible comparisons) with a chronic framing. A second version (assigned to about 30% of respondents) included 15 paired comparison questions with a temporary framing. A third version (about 30% of respondents) included five paired questions with the same chronic framing used in the first version and three population health equivalence questions. In the chronic framing of paired comparison questions, respondents were asked to imagine that each of the health states in the pair would last for the rest of a person's life. In the temporary framing of the paired comparison, respondents were asked to imagine that each health state would last for 1 week. All survey versions included questions on basic sociodemographic attributes (age, sex, educational level, income, and diseases that they have had). In randomly selected subsamples of respondents, the same pair of health states was presented in the first and 15th paired comparison questions to allow assessment of test–retest reliability.

Survey questionnaires were translated from English into Dutch, Hungarian, Italian, and Swedish by native speakers with a medical background, and subsequently translated back into English. The translations were verified independently by bilingual native speakers.

We report on estimated disability weights for the set of 235 states included in GBD 2013, consisting of the following categories: 135 states that were included in both GBD 2010 and the European disability weights measurement studies, with identical lay descriptions in both surveys; 50 states that are in GBD 2013 and were included in the GBD 2010 surveys and not in the European disability weights measurement study, but for which updated estimates emerge from the new analysis of pooled data; 30 states that were outcomes in both the GBD 2010 and 2013 studies, but for which GBD 2013 estimates are based on revised lay descriptions included in the European disability weights measurement study; 20 states that are new to GBD 2013, including 18 for which the new European disability weights measurement study provide data, and an additional two states for which the GBD 2010 surveys provide data. These last two states are monocular distance vision loss and headaches from drug use, which were included in the GBD 2010 surveys (and not the European disability weights measurement study), but not included as sequelae in the GBD 2010 burden of disease calculations.

Statistical analysis

We did all analyses in Stata/MP (version 11) and used the same analytic procedures applied in the GBD 2010 disability weights measurement study.[21] First, we pooled data for paired comparison responses from the original GBD 2010 dataset and the new European disability weights measurement study dataset. We ran probit regression analyses on the choice responses in the pooled data, with indicator variables for each health state that took the value 1 for the first state in a paired comparison, –1 for the second state in a paired comparison, and 0 for all states other than the pair being considered. The basic intuition of this modelling approach is that, in comparisons between states that produce similar losses of health, respondents are likely to disagree on which is healthier; in comparisons between states that produce very different health losses, respondents are more likely to agree on the ordering. The modelling strategy formalises this intuition into a statistical framework to infer the distances between values attached to different states on the basis of the observed frequencies of responses to paired comparison questions. With the probit regression, a second analytic step is needed to anchor the resulting estimates onto the 0–1 disability weights scale. In view of a separate analysis[29] that suggested that the population health equivalence questions did not produce valid responses in the European disability weights measurement study, we anchored results from the probit regression analysis onto the 0–1 scale using population health equivalence data from the GBD 2010 surveys and the same rescaling approach used in that study.[21] The approach relied on a linear regression of the probit coefficients from analysis of paired comparisons in the pooled dataset on the logit-transformed disability weight estimates derived from interval regression of the population health equivalence responses. We then estimated mean values for disability weights on the natural 0–1 scale using numerical integration. Analyses included a bootstrapping approach with 1000 replicate samples to estimate uncertainty about mean values.

In addition to the main analysis used to establish disability weights for GBD 2013, we ran a secondary analysis to examine one of the key design features of the study, which is the way that chronic and temporary outcomes were assessed to develop comparable weights for all sequelae in the study. In the design, to avoid conflation of severity with duration, all paired comparison questions in a given survey version were standardised for duration of health states—ie, they asked people to compare two outcomes that both lasted an entire lifetime or 1 week, depending on whether respondents were randomly allocated to the chronic or temporary version. However, because overlap existed, by design, between

subsets of health states that were included in the two versions of the survey, we analysed all paired comparison responses together, with the overlapping states bridging the two sets to enable this simultaneous analysis and estimation of weights on a single scale. Implicitly, this approach assumes that paired comparison responses will be indifferent to whether outcomes are chronic or temporary. To test this assumption, we ran separate analyses on the chronic and temporary responses and compared the resulting coefficients.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Table 1 shows the characteristics of the study population, combining the sample used in the GBD 2010 disability weights measurement study with that in the European disability weights measurement study. Overall, the combined sample consists of 30 230 respondents from the GBD 2010 surveys and 30 660 from the European disability weights measurement study. As expected, the new sample is characterised by high levels of educational attainment typical of European populations and similar to the distribution of educational levels in the USA-representative telephone survey. The European disability weights measurement study, by the design of the quota sampling scheme, shows an even balance by sex and a concentration of respondents in the age groups between 30 years and 65 years (summarised in table 1 in the categories 30–49 years and 50–69 years, but with no respondents aged 66–69 years, by design of the study). Examining test–retest reliability across both the GBD and European disability weights measurement studies, we noted that of 7810 participants who responded to the same pair of states in the first and 15th paired comparison questions, 5574 (71%) of these responses were consistent. Overall consistency was examined with heat maps of response probabilities, both for the original GBD surveys[21] and the European disability weights measurement study,[29] which substantiated a high degree of internal consistency in the aggregate response patterns to paired comparison questions.

Table 2 presents the estimated disability weights and 95% uncertainty intervals for all 235 states in GBD 2013. Across all states, estimated weights ranged from 0·003 (95% uncertainty interval 0·001–0·007) for mild distance vision impairment to 0·778 (0·606–0·900) for acute schizophrenia. The median weight across all states was 0·128 (IQR 0·040–0·291), with 59 (25%) of the states having disability weights of less than 0·04.

For the purposes of description of results here, we report separately on disability weights for several different categories of states, distinguished by whether the states appeared in the GBD 2010 survey, European disability weights measurement study, or both, and whether the GBD 2013 weights are based on original (GBD 2010) or modified lay descriptions. The 235 health states used in GBD 2013 include 185 that were also used in GBD 2010 (of 220 GBD 2010 states in total). Although only a subset of 135 of these states were included in the European disability weights measurement study, all disability weights have been re-estimated for the GBD 2013 on the basis of the full pooled dataset. For the subset of 135 included in both surveys, the revised weights for GBD 2013 take direct account of the additional paired comparison responses elicited in the new European disability weights measurement study; for the remaining 50, the estimated weights for GBD 2013 are affected indirectly by the new data because all weights are estimated simultaneously on the basis of the full information contained in the pooled dataset, comprising interlocked paired comparison responses across all categories of states.

The appendix (p 12–17) presents the disability weights in GBD 2013 alongside the GBD 2010 weights for the 185 states common to both studies. Comparing the two sets of weights (figure 1), we find a high degree of correlation overall (Pearson's r 0·992 [95% uncertainty interval 0·989–0·994]). For 36 (19%) states, the absolute value of the difference between the previous and new weights was less than 0·002; for 105 (57%) states, the absolute value of the difference was less than 0·01; the largest absolute value of the difference was 0·13 for

	Household surveys (n=10 579)	USA surveys (n=3323)	GBD 2010 web-based surveys (n=16 328)	European surveys (n=30 660)
Age (years)				
18–29	3416 (32%)	181 (5%)	5186 (32%)	6338 (21%)
30–49	4799 (45%)	870 (26%)	6660 (41%)	13 989 (46%)
50–69	2003 (19%)	1412 (42%)	4127 (25%)	10 333 (34%)
≥70	361 (3%)	852 (26%)	355 (2%)	0
Unknown	0	8 (<1%)	0	0
Sex				
Men	5814 (55%)	1230 (37%)	5268 (32%)	14 719 (48%)
Women	4765 (45%)	2092 (63%)	11 011 (67%)	15 941 (52%)
Unknown	0	11 (<1%)	49 (<1%)	0
Education				
None	2234 (21%)	8 (<1%)	59 (<1%)	144 (<1%)
Primary	2785 (26%)	75 (2%)	11 (<1%)	834 (3%)
Secondary	3362 (32%)	1162 (35%)	1035 (6%)	11 335 (37%)
Higher	2198 (21%)	2063 (62%)	15 173 (93%)	18 347 (60%)
Unknown	0	15 (<1%)	50 (<1%)	0

Data are n (%). GBD=Global Burden of Disease.

Table 1: **Characteristics of the study population**

	Estimate
Infectious disease	
Infectious disease	
Acute episode, mild	0·006 (0·002–0·012)
Acute episode, moderate	0·051 (0·032–0·074)
Acute episode, severe	0·133 (0·088–0·190)
Post-acute effects (fatigue, emotional lability, and insomnia)	0·219 (0·148–0·308)
Diarrhoea	
Mild	0·074 (0·049–0·104)
Moderate	0·188 (0·125–0·264)
Severe	0·247 (0·164–0·348)
Epididymo-orchitis	0·128 (0·086–0·180)
Herpes zoster	0·058 (0·035–0·090)
HIV: symptomatic, pre-AIDS	0·274 (0·184–0·377)
HIV/AIDS: receiving antiretroviral treatment	0·078 (0·052–0·111)
AIDS: not receiving antiretroviral treatment	0·582 (0·406–0·743)
Intestinal nematode infections: symptomatic	0·027 (0·015–0·043)
Lymphatic filariasis: symptomatic	0·109 (0·073–0·154)
Ear pain	0·013 (0·007–0·024)
Tuberculosis	
Without HIV infection	0·333 (0·224–0·454)
With HIV infection	0·408 (0·274–0·549)
Cancer	
Cancer	
Diagnosis and primary treatment	0·288 (0·193–0·399)
Metastatic	0·451 (0·307–0·600)
Mastectomy	0·036 (0·020–0·057)
Stoma	0·095 (0·063–0·131)
Terminal phase	
With medication (for cancers and end-stage kidney or liver disease)	0·540 (0·377–0·687)
Without medication (for cancers and end-stage kidney or liver disease)	0·569 (0·389–0·727)
Cardiovascular and circulatory disease	
Acute myocardial infarction	
Days 1–2	0·432 (0·288–0·579)
Days 3–28	0·074 (0·049–0·105)
Angina pectoris	
Mild	0·033 (0·020–0·052)
Moderate	0·080 (0·052–0·113)
Severe	0·167 (0·110–0·240)
Cardiac conduction disorders and cardiac dysrhythmias	0·224 (0·151–0·312)
Claudication	0·014 (0·007–0·025)
Heart failure	
Mild	0·041 (0·026–0·062)
Moderate	0·072 (0·047–0·103)
Severe	0·179 (0·122–0·251)
Stroke	
Long-term consequences, mild	0·019 (0·010–0·032)
Long-term consequences, moderate	0·070 (0·046–0·099)
Long-term consequences, moderate, plus cognition problems	0·316 (0·206–0·437)

(Table 2 continues in next column)

	Estimate
(Continued from previous column)	
Long-term consequences severe	0·552 (0·377–0·707)
Long-term consequences, severe, plus cognition problems	0·588 (0·411–0·744)
Diabetes and digestive and genitourinary disease	
Diabetic foot	0·020 (0·010–0·034)
Diabetic neuropathy	0·133 (0·089–0·187)
Chronic kidney disease (stage 4)	0·104 (0·070–0·147)
End-stage renal disease	
With kidney transplantation	0·024 (0·014–0·039)
On dialysis	0·571 (0·398–0·725)
Decompensated liver cirrhosis	0·178 (0·123–0·250)
Gastric bleeding	0·325 (0·209–0·462)
Crohn's disease or ulcerative colitis	0·231 (0·156–0·320)
Benign prostatic hypertrophy: symptomatic	0·067 (0·043–0·097)
Urinary incontinence	0·139 (0·094–0·198)
Stress incontinence	0·020 (0·011–0·035)
Impotence	0·017 (0·009–0·030)
Infertility	
Primary	0·008 (0·003–0·015)
Secondary	0·005 (0·002–0·011)
Chronic respiratory disease	
Asthma	
Controlled	0·015 (0·007–0·026)
Partly controlled	0·036 (0·022–0·055)
Uncontrolled	0·133 (0·086–0·192)
Chronic obstructive pulmonary disease and other chronic respiratory diseases	
Mild	0·019 (0·011–0·033)
Moderate	0·225 (0·153–0·310)
Severe	0·408 (0·273–0·556)
Neurological disorders	
Dementia	
Mild	0·069 (0·046–0·099)
Moderate	0·377 (0·252–0·508)
Severe	0·449 (0·304–0·595)
Headache	
Migraine	0·441 (0·294–0·588)
Tension-type	0·037 (0·022–0·057)
Medication overuse	0·223 (0·146–0·313)
Multiple sclerosis	
Mild	0·183 (0·124–0·253)
Moderate	0·463 (0·313–0·613)
Severe	0·719 (0·534–0·858)
Epilepsy	
Severe (seizures once per month or more)	0·552 (0·375–0·710)
Less severe (seizures less than once per month)	0·263 (0·173–0·367)
Parkinson's disease	
Mild	0·010 (0·005–0·019)
Moderate	0·267 (0·181–0·372)
Severe	0·575 (0·396–0·730)

(Table 2 continues in next column)

	Estimate
(Continued from previous column)	
Mental, behavioural, and substance use disorders	
Alcohol use disorder	
Very mild	0·123 (0·082–0·177)
Mild	0·235 (0·160–0·327)
Moderate	0·373 (0·248–0·508)
Severe	0·570 (0·396–0·732)
Fetal alcohol syndrome	
Mild	0·016 (0·008–0·030)
Moderate	0·056 (0·035–0·083)
Severe	0·179 (0·119–0·257)
Cannabis dependence	
Mild	0·039 (0·024–0·060)
Moderate to severe	0·266 (0·178–0·364)
Amphetamine dependence	
Mild	0·079 (0·051–0·114)
Moderate to severe	0·486 (0·329–0·637)
Cocaine dependence	
Mild	0·116 (0·074–0·165)
Moderate to severe	0·479 (0·324–0·634)
Heroin and other opioid dependence	
Mild	0·335 (0·221–0·473)
Moderate to severe	0·697 (0·510–0·843)
Anxiety disorders	
Mild	0·030 (0·018–0·046)
Moderate	0·133 (0·091–0·186)
Severe	0·523 (0·362–0·677)
Major depressive disorder	
Mild episode	0·145 (0·099–0·209)
Moderate episode	0·396 (0·267–0·531)
Severe episode	0·658 (0·477–0·807)
Bipolar disorder	
Manic episode	0·492 (0·341–0·646)
Residual state	0·032 (0·018–0·051)
Schizophrenia	
Acute state	0·778 (0·606–0·900)
Residual state	0·588 (0·411–0·754)
Anorexia nervosa	0·224 (0·150–0·312)
Bulimia nervosa	0·223 (0·149–0·311)
Attention-deficit hyperactivity disorder	0·045 (0·028–0·066)
Conduct disorder	0·241 (0·159–0·341)
Asperger's syndrome	0·104 (0·071–0·147)
Autism	0·262 (0·176–0·365)
Borderline intellectual functioning	0·011 (0·005–0·020)
Intellectual disability	
Mild	0·043 (0·026–0·064)
Moderate	0·100 (0·066–0·142)
Severe	0·160 (0·107–0·226)
Profound	0·200 (0·133–0·283)
Hearing and vision loss	
Hearing loss	
Mild	0·010 (0·004–0·019)

(Table 2 continues in next column)

	Estimate
(Continued from previous column)	
Moderate	0·027 (0·015–0·042)
Severe	0·158 (0·105–0·227)
Profound	0·204 (0·134–0·288)
Complete	0·215 (0·144–0·307)
Mild, with ringing	0·021 (0·012–0·036)
Moderate, with ringing	0·074 (0·049–0·107)
Severe, with ringing	0·261 (0·175–0·360)
Profound, with ringing	0·277 (0·182–0·387)
Complete, with ringing	0·316 (0·212–0·435)
Distance vision	
Mild impairment	0·003 (0·001–0·007)
Moderate impairment	0·031 (0·019–0·049)
Severe impairment	0·184 (0·125–0·258)
Blindness	0·187 (0·124–0·260)
Monocular impairment	0·017 (0·009–0·029)
Presbyopia	0·011 (0·005–0·020)
Musculoskeletal disorders	
Low back pain	
Mild	0·020 (0·011–0·035)
Moderate	0·054 (0·035–0·079)
Severe, without leg pain	0·272 (0·182–0·373)
Severe, with leg pain	0·325 (0·219–0·446)
Most severe, without leg pain	0·372 (0·250–0·506)
Most severe, with leg pain	0·384 (0·256–0·518)
Neck pain	
Mild	0·053 (0·034–0·078)
Moderate	0·114 (0·075–0·162)
Severe	0·229 (0·153–0·317)
Most severe	0·304 (0·202–0·415)
Musculoskeletal problems	
Legs, mild	0·023 (0·013–0·037)
Legs, moderate	0·079 (0·054–0·110)
Legs, severe	0·165 (0·112–0·232)
Arms, mild	0·028 (0·017–0·045)
Arms, moderate	0·117 (0·080–0·163)
Generalised, moderate	0·317 (0·216–0·440)
Generalised, severe	0·581 (0·403–0·739)
Gout: acute	0·295 (0·196–0·409)
Injury	
Amputation	
Finger(s), excluding thumb	0·005 (0·002–0·010)
Thumb: long term	0·011 (0·005–0·021)
One arm: long term, with treatment	0·039 (0·024–0·059)
One arm: long term, without treatment	0·118 (0·079–0·167)
Both arms: long term, with treatment	0·123 (0·081–0·176)
Both arms: long term, without treatment	0·383 (0·251–0·525)
Toe	0·006 (0·002–0·012)
One leg: long term, with treatment	0·039 (0·023–0·059)
One leg: long term, without treatment	0·173 (0·118–0·240)
Both legs: long term, with treatment	0·088 (0·057–0·124)

(Table 2 continues in next column)

	Estimate
(Continued from previous column)	
Both legs: long term, without treatment	0·443 (0·297–0·589)
Burns	
<20% of total surface area without lower airway burns: short term, with or without treatment	0·141 (0·094–0·196)
<20% of total surface area or <10% of total surface area if head or neck, or hands or wrist are involved: long term, with or without treatment	0·016 (0·008–0·028)
≥20% of total surface area: short term, with or without treatment	0·314 (0·211–0·441)
≥20% of total surface area or ≥10% of total surface area if head or neck, or hands or wrist are involved: long term, with treatment	0·135 (0·092–0·190)
≥20% of total surface area or ≥10% of total surface area if head or neck, or hands or wrist are involved: long term, without treatment	0·455 (0·302–0·601)
Lower airway: with or without treatment	0·376 (0·240–0·524)
Crush injury: short or long term, with or without treatment	0·132 (0·089–0·189)
Dislocation	
Hip: long term, with or without treatment	0·016 (0·008–0·028)
Knee: long term, with or without treatment	0·113 (0·075–0·160)
Shoulder: long term, with or without treatment	0·062 (0·041–0·088)
Other injuries of muscles and tendons (consisting of sprains, strains, and dislocations other than shoulder, knee, or hip)	0·008 (0·003–0·015)
Drowning and non-fatal submersion: short or long term, with or without treatment	0·247 (0·164–0·341)
Fracture	
Clavicle, scapula, or humerus: short or long term, with or without treatment	0·035 (0·021–0·053)
Face bone: short or long term, with or without treatment	0·067 (0·044–0·097)
Foot bones: short term, with or without treatment	0·026 (0·015–0·043)
Foot bones: long term, without treatment	0·026 (0·015–0·042)
Hand: short term, with or without treatment	0·010 (0·005–0·019)
Hand: long term, without treatment	0·014 (0·007–0·025)
Neck of femur: short term, with or without treatment	0·258 (0·172–0·356)
Neck of femur: long term, with treatment	0·058 (0·038–0·084)
Neck of femur: long term, without treatment	0·402 (0·269–0·541)
Other than neck of femur: short term, with or without treatment	0·111 (0·074–0·156)
Other than neck of femur: long term, without treatment	0·042 (0·027–0·063)
Patella, tibia or fibula, or ankle: short term, with or without treatment	0·050 (0·032–0·075)
Patella, tibia or fibula, or ankle: long term, with or without treatment	0·055 (0·036–0·081)
Pelvis: short term	0·279 (0·188–0·384)
Pelvis: long term	0·182 (0·123–0·253)
Radius or ulna: short term, with or without treatment	0·028 (0·016–0·046)

(Table 2 continues in next column)

	Estimate
(Continued from previous column)	
Radius or ulna: long term, without treatment	0·043 (0·028–0·064)
Skull: short or long term, with or without treatment	0·071 (0·048–0·100)
Sternum or one or two ribs: short term, with or without treatment	0·103 (0·068–0·145)
Vertebral column: short or long term, with or without treatment	0·111 (0·075–0·156)
Treated, long term	0·005 (0·002–0·010)
Injured nerves	
Short term	0·100 (0·067–0·140)
Long term	0·113 (0·076–0·157)
Injury to eyes: short term	0·054 (0·035–0·081)
Concussion	0·110 (0·074–0·158)
Traumatic brain injury	
Severe: short term, with or without treatment	0·214 (0·141–0·297)
Long-term consequences, minor, with or without treatment	0·094 (0·063–0·133)
Long-term consequences, moderate, with or without treatment	0·231 (0·156–0·324)
Long-term consequences, severe, with or without treatment	0·637 (0·462–0·789)
Open wound: short term, with or without treatment	0·006 (0·002–0·012)
Poisoning: short term, with or without treatment	0·163 (0·109–0·227)
Severe chest injury	
Long term, with or without treatment	0·047 (0·030–0·070)
Short term, with or without treatment	0·369 (0·248–0·501)
Spinal cord lesion	
Below neck: treated	0·296 (0·198–0·414)
Below neck: untreated	0·623 (0·434–0·777)
At neck: treated	0·589 (0·415–0·748)
At neck: untreated	0·732 (0·544–0·871)
Other	
Abdominopelvic problem	
Mild	0·011 (0·005–0·021)
Moderate	0·114 (0·078–0·159)
Severe	0·324 (0·220–0·442)
Anaemia	
Mild	0·004 (0·001–0·008)
Moderate	0·052 (0·034–0·076)
Severe	0·149 (0·101–0·209)
Periodontitis	0·007 (0·003–0·014)
Dental caries: symptomatic	0·010 (0·005–0·019)
Severe tooth loss	0·067 (0·045–0·095)
Disfigurement	
Level 1	0·011 (0·005–0·021)
Level 2	0·067 (0·044–0·096)
Level 3	0·405 (0·275–0·546)
Level 1, with itch or pain	0·027 (0·015–0·042)
Level 2, with itch or pain	0·188 (0·125–0·267)
Level 3, with itch or pain	0·576 (0·401–0·731)

(Table 2 continues in next column)

	Estimate
(Continued from previous column)	
Generic uncomplicated disease	
Worry and daily medication	0·049 (0·031–0·072)
Anxiety about diagnosis	0·012 (0·006–0·023)
Iodine-deficiency goitre	0·199 (0·133–0·276)
Kwashiorkor	0·051 (0·031–0·079)
Severe wasting	0·128 (0·082–0·183)
Speech problems	0·051 (0·032–0·078)
Motor impairment	
Mild	0·010 (0·005–0·019)
Moderate	0·061 (0·040–0·089)
Severe	0·402 (0·268–0·545)
Motor and cognitive impairments	
Mild	0·031 (0·018–0·050)
Moderate	0·203 (0·134–0·290)
Severe	0·542 (0·374–0·702)
Rectovaginal fistula	0·501 (0·339–0·657)
Vesicovaginal fistula	0·342 (0·227–0·478)
Thrombocytopenic purpura	0·159 (0·106–0·226)
Hypothyroidism	0·019 (0·010–0·032)
Hyperthyroidism	0·145 (0·096–0·202)
Vertigo	0·113 (0·074–0·158)

Data in parentheses are 95% uncertainty intervals.

Table 2: **Disability weights for 235 unique health states in the Global Burden of Disease 2013 study**

moderate-to-severe amphetamine dependence (for which the weight increased from 0·353 [95% uncertainty interval 0·215–0·525] in GBD 2010 to 0·486 [0·329–0·637] in GBD 2013). Differences were less than 5% for 68 (37%) states and less than 20% for 148 (80%) states; the largest relative difference was 61% for short-term fracture of the hand (for which the weight decreased from 0·025 [0·013–0·043] in GBD 2010 to 0·010 [0·005–0·019] in GBD 2013). Other states with large differences included severe infectious disease episodes and post-infection symptoms; treated HIV/AIDS cases; cardiac conduction disorders; diabetic neuropathy; controlled asthma; various substance use disorders; some outcomes relating to burns; several injury-related states, including short-term fractures and eye injuries; and a generic state consisting of worry and daily drugs for an uncomplicated disease (appendix). Within the subset of 50 states that are common to GBD 2010 and 2013 but not included in the European disability weights measurement study (appendix p 16–17), differences were substantially smaller than for those that were included in the European disability weights measurement study, as expected, with the maximum absolute difference across states being 0·02 and the maximum relative difference 18%.

For 30 states in GBD 2013, new descriptions were developed to replace those used in GBD 2010. The appendix (p 18–21) shows a comparison of weights for states that have had their lay descriptions revised for the 2013 study. Minor changes in lay descriptions produced the expected minor shifts in disability weights for some outcomes, including mild depression, motor and cognitive impairments, and intellectual disabilities. Revisions to a series of states relating to amputations resulted in a mixture of small and moderate changes to the associated disability weights. The most sizeable changes to disability weights were in those relating to hearing loss, for which addition of the previously omitted references to difficulties with communication and relating to other people resulted in increases in the disability weights for the more severe levels of hearing loss, with complete hearing loss changing from 0·033 (95% uncertainty interval 0·020–0·052) in the previously published GBD 2010 disability weights measurement study[21] to 0·215 (0·144–0·307) in GBD 2013. Another set of outcomes with substantial changes to the weights resulting from revision of lay descriptions was the series of four states relating to spinal cord lesions; the largest of these changes pertained to the two treated states, with the weight for treated spinal cord lesions below the neck increasing from 0·047 (0·028–0·072) to 0·296 (0·198–0·414), and for treated spinal cord lesions at neck level increasing from 0·369 (0·243–0·513) to 0·589 (0·415–0·748).

For a final grouping of outcomes new to GBD 2013, the appendix (p 22) presents lay descriptions and estimated disability weights. Of these disability weights, the lowest were for borderline intellectual functioning, at 0·011 (0·005–0·020), and distance vision, monocular impairment, at 0·017 (0·009–0·029). The highest disability weights were for severe epilepsy, at 0·552 (0·375–0·710), and mild heroin and other opioid dependence, at 0·335 (0·221–0·473).

To assess the implicit assumption that paired comparison responses are insensitive to whether outcomes are presented as either chronic or temporary (standardised for both outcomes within a pair, either as lifetime or 1 week duration, depending on the survey version), we ran separate probit analyses on paired comparison responses for states included in the survey version using a chronic formulation of the paired comparison (including 201 GBD 2013 states) and a temporary formulation (including 142 GBD 2013 states). For the 108 GBD 2013 states appearing in both sets, the coefficients from the separate probit regressions were highly correlated (Pearson's r 0·981 [95% uncertainty interval 0·973–0·987]). Figure 2 is a scatter plot of the results from the chronic and temporary regressions. The states with the largest deviations from the linear relation between the two sets of regression coefficients included speech problems and rectovaginal fistulas, which had chronic values that were less severe than what would be predicted from the temporary values, and untreated HIV/AIDS cases, which had a chronic value that was more severe than that expected in relation to the temporary value.

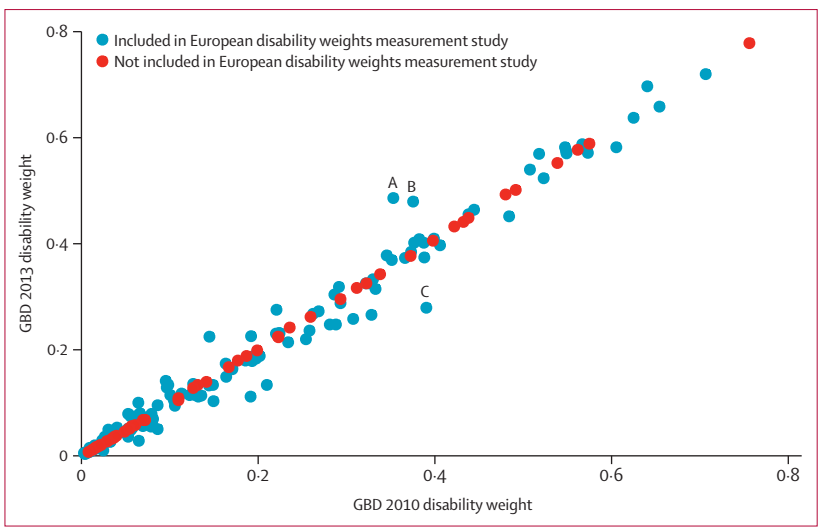

Figure 1: **Comparison of disability weights from GBD 2010 and 2013 for 135 states included and 50 not included in the European disability weights measurement study**
The three states with the largest absolute differences between the GBD 2010 and 2013 weights are labelled as follows: (A) amphetamine dependence: moderate to severe; (B) cocaine dependence: moderate to severe; (C) fracture of pelvis: short term. GBD=Global Burden of Disease.

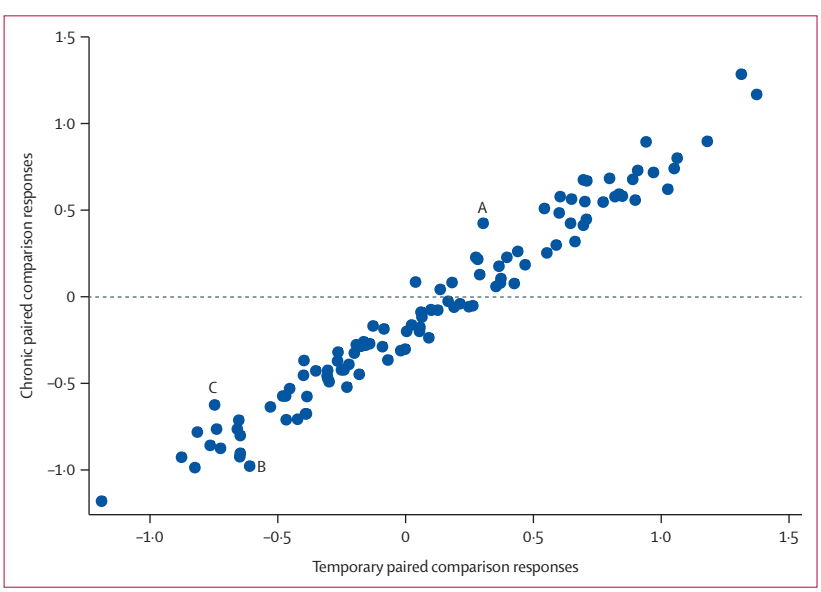

Figure 2: **Comparison of probit coefficients estimated from responses to the version of the questionnaire with the chronic framing versus the version of the questionnaire with the temporary framing**
The three states showing the largest deviations from the linear relation between the two sets of regression coefficients are labelled as follows: (A) speech problems; (B) AIDS cases not receiving antiretroviral treatment; (C) rectovaginal fistula.

Discussion

In this study, we build on the previous disability weights measurement study for GBD 2010, with a substantial expansion of the empirical database of comparative assessments for a range of health outcomes. In the new analysis, with a combined sample that doubles the number of respondents from the original dataset, we note that most disability weights are minimally affected by the change in composition of the sample towards one in which respondents from European countries now make up most survey participants. Several exceptions to this general finding exist, however, including some states for which weights are substantially higher in the European disability weights measurement study than in GBD 2010 (eg, treated HIV/AIDS cases, cardiac conduction disorders, controlled asthma, short-term burn injuries, and a generic state for worry and daily drug use associated with an uncomplicated diagnosed disease), and some for which they are substantially lower in the new study than in the old study (eg, severe infections and short-term states relating to fractures of the hands, arms, or legs). Although this study was not designed to develop or test specific hypotheses about why the introduction of a large European respondent sample might have produced results that differed from existing estimates in these ways, such questions are worth further exploration, especially in cases for which these differences could have substantial implications for the magnitude or ranking of particular disease burdens.

With incorporation of the new survey data from four European countries in estimation of disability weights for GBD 2013, the cumulative evidence base for measurement of these weights is now mostly made up of respondents from high-income settings. The previously noted finding that paired comparison responses are largely consistent across very distinct settings[21] and the finding in this study that most disability weights did not change substantially in the new pooled analysis together offer some reassurance that most disease burden calculations will be reasonably robust to this shift. However, the exceptions noted above of some disability weights that did change substantially in the new analysis are important to acknowledge and their implications considered. Efforts to collect more data for disability weights in different settings are planned or underway. Work should continue on assessment of the sensitivity of key GBD findings to changes in the composition of respondents who inform estimation of disability weights for the worldwide analysis.

In addition to examination of the sensitivity of disability weights to the composition of sample respondents, this study also offers evidence for the effect of inclusion of specific elements in or exclusion of the elements from lay descriptions of outcomes presented as stimuli for paired comparison questions. Our analysis has revealed several states for which the changes in lay descriptions from those used in GBD 2010 were associated with major revisions of the weights, notably the increase in weights for a range of hearing loss outcomes, especially at severe levels, when explicitly accounting for effects of social isolation. These changes reinforce the inherent limitation in an approach based on use of brief lay descriptions in presentation of outcomes in paired comparisons, which is that, in some cases, responses are evidently highly sensitive to particular details in these descriptions. An

example of a potential concern with some lay descriptions, which persists from the original GBD 2010 study, is the case of illicit drug use disorders, which have weights that are higher than those reported in other studies. We have hypothesised that part of the explanation for this result is that lay descriptions for illicit drug use disorders include explicit reference to drug use as the cause of the functional outcomes described, whereas we have deliberately avoided this sort of labelling in other health states. Further examination of the potential for bias relating to explicit labelling of causes such as drug use disorders remains a priority. Close inspection of these results, as with the original publication of the disability weights for GBD 2010, will undoubtedly lead to identification of other examples of weights that seem counterintuitive. These types of findings need to be addressed with empirical investigation to understand whether the weights in question are sensitive to specific elements in the lay descriptions, or whether some other explanation might emerge.

A major methodological and empirical agenda remains to improve the basis of standardised description of outcomes, which depends on two key advances. First, methods are needed to parsimoniously present the major functional and symptom-related attributes for a highly diverse set of health outcomes captured in the GBD study, which has to balance the diversity of attributes with the cognitive demands of people comparing outcomes that differ in many dimensions. Second, these descriptions need to be driven by empirical data for the profiles of individuals living with each outcome, drawn from representative samples of such individuals; these samples will, in many cases, be difficult to identify or recruit, and this would represent a vast and ambitious new data collection enterprise.

One of the contributions of this study is new evidence for the robustness of paired comparison responses to framing of outcomes in the paired comparison as either chronic or temporary. We note that health state comparisons are largely insensitive to this framing choice. This finding suggests that our approach of simultaneous estimation of cardinal severity values from a pooled dataset with a combination of responses to chronic and temporary paired comparisons is a reasonable methodological choice.

On the other hand, we reiterate some important limitations in the study that carry over from the previous disability weights measurement study for GBD 2010. Our approach to anchor disability weights on a 0–1 scale continues to depend on a small number of responses to a population health equivalence question that was included in the GBD 2010 web-based surveys. Elsewhere,[29] we have noted that the discrete choice formulation of this question seems to be susceptible to a large amount of measurement error in a general population sample (as used in the European disability weights measurement study, by contrast with the self-selected sample of respondents to the GBD 2010 web-based survey), making this type of question format unsuitable for use as an anchoring device. Specifically, we examined the European population health equivalence responses for both sensitivity to scope and responsiveness to variation in severity of the different outcomes under consideration. In both cases, patterns of response were consistent with large proportions of respondents answering randomly, which compromised the goal of scaling disability weights on the basis of the results of these responses. Moving forward, new approaches might be needed to anchor estimates from paired comparisons onto the scale needed for disability weights. At present, limited evidence exists for how trade-offs between non-fatal outcomes and changes in mortality vary across settings, so resolution of this question continues to await further empirical examination. These questions, both methodological and empirical, remain important priorities for investigation within the continuing GBD research agenda.

As the GBD enterprise has committed to annual updates of the systematic reviews and pooling of available data for deaths, prevalence of disease, and the contribution of risk factors, aiming to take account of all the available evidence, in the same spirit, new information about disability weights will be reviewed and incorporated into GBD estimates as it becomes available. The high degree of consistency between the findings of the original GBD and the recent European disability weights measurement study lends further support to their pooled use to inform comparable burden of disease estimates for all countries. Evidently, more important than between-country variation are the substantial changes in estimated levels of disability weights that can result from addition of a new symptom or alteration of the wording of lay descriptions of a health state. As the GBD study expands, opportunities for new survey work will need to be seized to further explore some of these methodological issues and allow consistent valuation of health states characterising newly added diseases.

Contributors

JAS led the conception and design of the GBD 2010 disability weights measurement study, development of data collection instruments and protocols, and analysis and interpretation of data, and wrote the first draft of this report. For the European disability weights measurement study, JAH, JAS, and TV led the conception and design of the study; JAH, JAS, TV, SP, AC, MK, AHH, and CJLM developed the study design; JAS, JAH, TV, and AD developed survey instruments and selected and developed health state descriptions; JAH supervised data collection; and CMdN, JAS, JAH, BD, and NS analysed and interpreted data. All authors critically reviewed the report and approved the final version.

Declaration of interests

AC is a staff member at the European Centre for Disease Prevention and Control, which partly funded the European disability weights measurement study. All other authors declare no competing interests.

Acknowledgments

The European disability weights measurement study, which supplied part of the dataset analysed in this report, was supported by funding from the European Centre for Disease Prevention and Control and Bill and Melinda Gates Foundation.

References

1 Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bull World Health Organ* 1994; **72:** 429–45.

2 Murray CJ. Rethinking DALYs. In: Murray CJ, Lopez AD, eds. 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: Harvard School of Public Health, 1996.

3 Salomon JA, Mathers CD, Chatterji S, Sadana R, Ustun TB, Murray CJ. Quantifying individual levels of health: definitions, concepts and measurement issues. In: Murray CJ, Evans DB, eds. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization, 2003: 301.

4 Salomon JA, Murray CJ, Ustun TB, Chatterji S. Health state valuations in summary measures of population health. In: Murray CJ, Evans DB, eds. Health systems performance assessment: debate, methods, and empiricism. Geneva: World Health Organization, 2003: 409–36.

5 Arnesen T, Nord E. The value of DALY life: problems with ethics and validity of disability adjusted life years. *BMJ* 1999; **319:** 1423–25.

6 Ustun TB, Rehm J, Chatterji S, et al. Multiple-informant ranking of the disabling effects of different health conditions in 14 countries. WHO/NIH Joint Project CAR Study Group. *Lancet* 1999; **354:** 111–15.

7 Ustun TB, Saxena S, Rehm J, Bickenbach J. Are disability weights universal? WHO/NIH Joint Project CAR Study Group. *Lancet* 1999; **354:** 1306.

8 Groce NE. Disability in cross-cultural perspective: rethinking disability. *Lancet* 1999; **354:** 756–57.

9 Jelsma J, Chivaura VG, Mhundwa K, De Weerdt W, de Cock P. The global burden of disease disability weights. *Lancet* 2000; **355:** 2079–80.

10 AbouZahr C, Vaughan JP. Assessing the burden of sexual and reproductive ill-health: questions regarding the use of disability-adjusted life years. *Bull World Health Organ* 2000; **78:** 655–66.

11 Bastian H. A consumer trip into the world of the DALY calculations: an Alice-in-Wonderland experience. *Reprod Health Matters* 2000; **8:** 113–16.

12 Metts RL. The fatal flaw in the disability adjusted life year. *Disabil Soc* 2001; **16:** 449–52.

13 Jelsma J, De Weerdt W, De Cock P. Disability adjusted life years (DALYs) and rehabilitation. *Disabil Rehabil* 2002; **24:** 378–82.

14 Mansley EC, Elbasha EH. Preferences and person trade-offs: forcing consistency or inconsistency in health-related quality of life measures? *Health Econ* 2003; **12:** 187–98.

15 Allotey P, Reidpath D, Kouame A, Cummins R. The DALY, context and the determinants of the severity of disease: an exploratory comparison of paraplegia in Australia and Cameroon. *Soc Sci Med* 2003; **57:** 949–58.

16 Reidpath DD, Allotey PA, Kouame A, Cummins RA. Measuring health in a vacuum: examining the disability weight of the DALY. *Health Policy Plan* 2003; **18:** 351–56.

17 Mont D. Measuring health and disability. *Lancet* 2007; **369:** 1658–63.

18 King CH, Bertino AM. Asymmetries of poverty: why global burden of disease valuations underestimate the burden of neglected tropical diseases. *PLoS Negl Trop Dis* 2008; **2:** e209.

19 Rehm J, Frick U. Valuation of health states in the US study to establish disability weights: lessons from the literature. *Int J Methods Psychiatr Res* 2010; **19:** 18–33.

20 Salomon JA. New disability weights for the global burden of disease. *Bull World Health Organ* 2010; **88:** 879.

21 Salomon JA, Vos T, Hogan DR, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. *Lancet* 2012; **380:** 2129–43.

22 Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380:** 2163–96.

23 Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380:** 2197–223.

24 Salomon JA, Wang H, Freeman MK, et al. Healthy life expectancy for 187 countries, 1990–2010: a systematic analysis for the Global Burden Disease Study 2010. *Lancet* 2012; **380:** 2144–62.

25 Nord E. Disability weights in the Global Burden of Disease 2010: unclear meaning and overstatement of international agreement. *Health Policy* 2013; **111:** 99–104.

26 Taylor HR, Jonas JB, Keeffe J, et al. Disability weights for vision disorders in Global Burden of Disease study. *Lancet* 2013; **381:** 23.

27 Voigt K, King NB. Disability weights in the global burden of disease 2010 study: two steps forward, one step back? *Bull World Health Organ* 2014; **92:** 226–28.

28 Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015; **386:** 743–800.

29 Haagsma JA, Maertens de Noordhout C, Polinder S, et al. Assessing disability weights based on the responses of 30,660 people from four European countries. *Popul Health Metr* 2015; **13:** 10.

30 Kretzschmar M, Mangen MJ, Pinheiro P, et al. New methodology for estimating the burden of infectious diseases in Europe. *PLoS Med* 2012; **9:** e1001205.