

Disability weights in the global burden of disease 2010 study: two steps forward, one step back?

Kristin Voigt^a & Nicholas B King^b

In the global burden of disease (GBD) project, the disability-adjusted life year (DALY) is used as a measure of disease burden. DALYs combine years of life lost to premature death and years lived with a disability or health condition, weighted for the severity of the condition. This combined measure makes it possible to compare health conditions with different symptoms and outcomes using a standardized metric.

Initiated in the early 1990s, the methods used to measure the GBD recently underwent major revision. Details about the new methods and up-to-date information about the state of health in the world were published in the December 2012 issue of the *Lancet*, which was devoted entirely to the GBD 2010 study.¹ Although the new methods used to calculate the GBD are a considerable improvement over previous ones, central concerns about the utility of DALYs for making decisions about the prioritization of health conditions and the allocation of resources remain unaddressed.

General concerns and how they were addressed

Since their inception, GBD studies have been criticized for their methodological and normative choices in four areas: the use of age weighting (i.e. the assignment of different weights to years of life lost at different ages); the use of different life expectancies for men and women; discounting (i.e. the assignment of a lower weight to years of life lost in the future); and the determination of disability weights, which are intended to capture the severity of a condition.

The team that conducted the GBD 2010 study took these criticisms seriously and enlisted experts in multiple disciplines to help revise its methods. In response to the first three criticisms it

dropped discounting, made age weights uniform and used the same life expectancy for men and women (86 years) to calculate years of life lost.¹ To address the fourth criticism, the team revised its method for determining disability weights. These are standardized values that are assigned to non-fatal health outcomes to capture their severity on a scale between 0 (full health) and 1 (death). For example, in the GBD 2010 study, complete hearing loss and severe chronic neck pain were assigned disability weights of 0.033 and 0.286, respectively.

Many critics of earlier GBD studies, which were originally commissioned in 1990 and have been updated twice since, were dissatisfied with the particular person trade-off method used to determine disability weights. The method consisted of asking respondents to choose between different hypothetical public health interventions.² Subsequently, critics questioned the feasibility of developing a standardized health measure. They argued, furthermore, that the disability weights incorporated into the DALY, which were originally elicited from a small group of highly educated health experts, could not be generalized to respondents from different geographic, cultural or socioeconomic backgrounds. Many doubted that a single universal measure of health could be validly applied to all settings, given marked differences among countries in baseline levels of health and socioeconomic development. Finally, some critics contended that “health” cannot – or should not – be separated from general welfare, which is shaped not only by an individual’s symptoms but also by the interaction of those symptoms with the environment.

The GBD 2010 study team formulated several responses to these concerns. First, the GBD 2010 study surveys employed brief lay descriptions of the symptoms and functional consequences

of various health states in an attempt to isolate losses in health from losses in welfare and from the impact of the social environment.³ This method reflected the assumption that the GBD study should capture “health” in a narrow sense. The insistence on a narrow understanding of health stemmed from the conviction that, even though the experience of a particular condition (e.g. paraplegia) can differ vastly depending on where a person lives and the level of care and social support available, the underlying health state remains the same. To support this contention, the GBD team conducted studies demonstrating broad consistency – with some exceptions regarding HIV infection and physical disabilities – in how health states are evaluated by respondents from different cultural backgrounds.⁴

Second, disability weights in the GBD 2010 study were based on data from household surveys conducted in different countries – Bangladesh, Indonesia, Peru, the United Republic of Tanzania and the United States of America – as well as from a web-based survey. Thus, the disability weights reflected responses obtained in different settings and demographic groups, including individuals with little formal education.³

Third, health state valuation in the GBD 2010 study was based on pairwise comparisons of different health states rather than person trade-offs. For the pairwise comparisons, respondents were asked to indicate which of two individuals in different states of health they considered “healthier overall”.^{3,5} This means that respondents were asked to focus on how “health” alone – rather than particular conditions in interaction with contextual factors – affected welfare or well-being.

Remaining concerns

The GBD 2010 team responded to the criticisms levelled at prior GBD studies

^a Institute for Health and Social Policy and Department of Philosophy, McGill University, 1130 Pine Ave West, Montreal, QC H3A 1A3, Canada.

^b Biomedical Ethics Unit, McGill University, Montreal, Canada.

Correspondence to Kristin Voigt (e-mail: kristin.voigt@mcgill.ca).

(Submitted: 21 June 2013 – Revised version received: 28 November 2013 – Accepted: 3 December 2013 – Published online: 10 January 2014)

in a transparent and comprehensive manner. Nevertheless, we believe that many of the criticisms remain valid, particularly when it comes to using GBD study results to inform decisions about resource allocation and disease control priorities.

First, despite claims that the methods followed in the GBD 2010 study isolated “health” from welfare, well-being and social context, a close reading of those methods illustrates how difficult it is to maintain a strict separation between “health” *simpliciter* and the effects of particular conditions on individuals in particular contexts. For example, the descriptions of disfigurement include the reactions of others to the disfigurement (“makes others feel uncomfortable”) and the implications of these reactions for the individual with the disfigurement (“causes the individual to avoid social contact”).⁵ Since disfigurement does not necessarily cause pain or other physical symptoms, it is “bad” primarily because of its adverse impact on social engagement. However, such impact is exactly what should be excluded if the goal is to capture “health” rather than health in interaction with an individual’s social context.

The GBD study is also inconsistent insofar as it mentions the social implications of some health states but not others. For example, it mentions the social consequences of disfigurement but not of human immunodeficiency virus (HIV) infection, which can be quite serious in many countries. It could be argued, for example, that many other health conditions besides disfigurement “make others feel uncomfortable” or “cause the individual to avoid social contact”. Similarly, only the descriptions of intellectual disabilities mention the condition’s impact on the ability to raise children without assistance – which certainly seems to be an impact on welfare rather than on health per se – despite the fact that many other conditions have similar implications. As the GBD team notes, differences in disability weights may be sensitive to this sort of variation.³ Future revisions of the GBD study methods should attend to these inconsistencies more closely.

Second, the revised methods do not warrant the conclusion that health valuations are universal. Salomon et al. find significant correlations across respondents from different backgrounds, which

they consider “compelling evidence that contradicts the prevailing hypothesis that assessments of disability must vary widely across samples with diverse cultural, educational, environmental, or demographic circumstances”.³ Although its methods are a vast improvement over previous ones, the GBD 2010 study is based on an idiosyncratic sample that is – as Salomon et al. acknowledge³ – not globally representative. For the household and web surveys combined, 35% of the respondents were from the United States, 64% were highly educated and only 17% had a primary education or less. Only four countries – Bangladesh, Indonesia, Peru and the United Republic of Tanzania – represent the multiplicity of cultures existing outside the United States and the entire sample is heavily skewed towards the highly educated. One may thus reasonably say that although the GBD 2010 study reflects a high level of agreement, this agreement pertains to a narrow construct of “health” within a rather narrow slice of humanity. In future, GBD study methods should be revised to target or oversample individuals with lower education and from a wider range of countries.

More importantly, this approach involves a certain amount of circular reasoning. The new health state valuation method employed for the GBD 2010 study was designed to exclude variations introduced by considerations of context, not to determine whether these considerations are relevant or vary across different respondents. Asking respondents to make “healthier than” comparisons abstracts from contextual factors to a greater extent than, for example, the person trade-off approach used in the previous version of the GBD study. The consistency in the responses to this narrower question is therefore not surprising. Although such agreement is reassuring, it does not invalidate concerns surrounding the exclusion of contextual factors from measures of health; it simply sidesteps them. The concern is not that agreement on some narrow conception of health is not possible, but that the narrow conception may not be useful for decisions about disease prioritization and resource allocation. A more appropriate response to this line of criticism would be to emphasize the distinction between using metrics like the DALY for monitoring

global health, and using them to determine priorities and allocate resources among different diseases in different countries. The DALY’s narrow focus and insensitivity to contextual factors make it useful for monitoring global health, but these characteristics also limit its utility for assessing interventions and establishing priorities.

For example, let us consider two countries with equal numbers of people with impaired vision. In one country corrective lenses are readily available and people with visual impairment have good job prospects; in the other, there are no corrective lenses and employment opportunities are few. In an abstract sense the amount of impaired vision is the same in both populations, but the impact of the impairment is radically different – a consideration that is vital to the distribution of resources. The important question is whether resource distribution should be based on an abstract understanding of the “health” effects of different conditions, or on the actual impact of the conditions on the welfare of individuals and populations. In this respect, the GBD 2010 study provides an indicator – the DALY – that is useful for disease surveillance but of limited use for establishing policy priorities and distributing resources. These limitations, we believe, must be spelt out explicitly if GBD data are to be a helpful and appropriate resource for guiding global health policy. The danger is not that the DALY is an invalid measure, but rather, that individuals and organizations may use it for purposes for which it is not suited.

Conclusion

The GBD 2010 study has undergone methodological improvements intended to address important objections that had been levelled against it. Such improvements have lent greater legitimacy to the project and to the use of its results for disease surveillance. The study offers policymakers a valuable resource for assessing the overall GBD. However, even this much-improved revision has not yet settled the debate over whether ostensibly universal disability weights are possible, desirable or useful for policy purposes. These questions carry less weight than many critics assume when GBD data are used to monitor global health. However, they must be

addressed if GBD data are to be used – as has been envisaged by researchers within and outside the GBD project – to inform decisions about health policy, resource allocation and health priorities.⁶ Ultimately, when it comes to deciding which conditions to prioritize or how to

best distribute health resources within and between countries, data from the GBD 2010 study must be supplemented with additional information regarding the impact of different conditions on the health and welfare of individuals in different locales. ■

Funding: This work was supported by the Canadian Institutes for Health Research (CIHR), Operating Grant EPP-122908, “Measuring Global Health”.

Competing interests: None declared.

References

1. Murray CJL, Ezzati M, Flaxman AD, Lim S, Lozano R, Michaud C et al. GBD 2010: design, definitions, and metrics. *Lancet* 2012;380:2063–6. doi: [http://dx.doi.org/10.1016/S0140-6736\(12\)61899-6](http://dx.doi.org/10.1016/S0140-6736(12)61899-6) PMID:23245602
2. Arnesen T, Nord E. The value of DALY life: problems with ethics and validity of disability adjusted life years. *BMJ* 1999;319:1423–5. doi: <http://dx.doi.org/10.1136/bmj.319.7222.1423> PMID:10574867
3. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A 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. doi: [http://dx.doi.org/10.1016/S0140-6736\(12\)61680-8](http://dx.doi.org/10.1016/S0140-6736(12)61680-8) PMID:23245605
4. Ustün TB, Rehm J, Chatterji S, Saxena S, Trotter R, Room R et al.; WHO/NIH Joint Project CAR Study Group. Multiple-informant ranking of the disabling effects of different health conditions in 14 countries. *Lancet* 1999;354:111–5. doi: [http://dx.doi.org/10.1016/S0140-6736\(98\)07507-2](http://dx.doi.org/10.1016/S0140-6736(98)07507-2) PMID:10408486
5. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010 [supplementary appendix]. *Lancet* 2012;380:2129–43. doi: [http://dx.doi.org/10.1016/S0140-6736\(12\)61680-8](http://dx.doi.org/10.1016/S0140-6736(12)61680-8) PMID:23245605
6. Horton R. GBD 2010: understanding disease, injury, and risk. *Lancet* 2012;380:2053–4. doi: [http://dx.doi.org/10.1016/S0140-6736\(12\)62133-3](http://dx.doi.org/10.1016/S0140-6736(12)62133-3) PMID:23245595

Corrigendum

In Volume 91, Issue 11, June 2013, page 893, an arrow should be added from “Policies to address inflows and outflows” to “Abroad”, and the arrow should be deleted from “Policies on production” to “Abroad”.

