Metrics in Global Health: Situated differences in the valuation of human life

Title in German:

Tiago Moreira, Department of Sociology, Durham University, Durham, United Kingdom.

Abstract

This paper explores the role of knowledge, standards and metrics in global health. Our point of departure is the observation that the emergence of ‘global health’ as a domain of research, policy and practice in the last three decades or so has coincided with an increased interest in the validation and use of measures of health such as the Disability Adjusted Life Year (DALY) in monitoring and assessing health equity across territories and populations. This ‘elective affinity’ between global health and health metrics has become the focus of scholarly debate in the social sciences.

In the paper, we seek to contextualise and critically discuss the different positions in this debate. We suggest that emplacing health metrics within the neo-liberal logic of health production –one where the ‘mechanisms of life’ are aligned with the maximisation of economic productivity- does not fully capture the interactive relationship between health measurement and the politics of health. We instead argue that this relationship has been characterised by controversy and uncertainty about how to interlock normative ideals and approaches to knowledge-making about health.

Reference address:
Escuela de Ciencias Humanas, Universidad del Rosario, Calle 12c #6-25. Bogotá, Colombia.

email:

Keywords: Metrics, DALY, Global Health, quantification, valuation

Introduction

This paper explores the role of knowledge, standards and metrics in global health. Our point of departure is the observation that the emergence of ‘global health’ as a domain of research, policy and practice in the last three decades or so has coincided with an increased interest in the validation and use of measures of health such as the Disability Adjusted Life Year
in monitoring and assessing health equity across territories and populations. This ‘elective affinity’ between global health and health metrics has become the focus of scholarly debate in the social sciences. In the paper, we first seek to contextualise and critically discuss the different positions in this debate. We suggest that emplacing health metrics within the neo-liberal logic of health production – one where the ‘mechanisms of life’ are aligned with the maximisation of economic productivity (Foucault, 2003) - does not fully capture the interactive relationship between health measurement and the politics of health. We instead argue that this relationship has been characterised by controversy and uncertainty about how to interlock normative ideals and approaches to knowledge-making about health.

This analysis results from approaching the subject from the perspective of science and technology studies (STS). STS investigates the complex, controversial processes through which knowledge and technology are brought together and to bear within social and culturally relevant settings and relation. Contrary to what is usually assumed, controversies – particularly public disputes of a technical variety hinging on a matter that concerns most people in society- are not mere surface phenomena, denoting only the biases of the opposing sides. Instead, controversies should be seen as windows to understand society-in-the-making. As Latour put it, controversies are “not simply a nuisance to be kept at bay, but what allows the social to be established” (Latour, 2005: 25). This is because they substantiate ways of knowing and acting on the world, and concurrently mould institutions, policies and programmes.

Controversies about standards and forms of measurement are particularly important from a sociological point of view. Standards – in our case, standardised forms of measuring and quantifying health – are explicit, formalised rules or specifications informing collective engagement with objects or persons in a particular realm of action; in this, they are able to ‘regulate and calibrate social life by rendering the … world equivalent across cultures, time, and geography’ (Timmermans and Epstein 2010: 70). The normative bearing of the work of calibrating and drawing equivalences is exquisitely encapsulated in Thevenot’s (1984) concept of ‘investment in form’, to refer to the process of making of durable conventional arrangements - ‘a stable relation with a certain lifespan’- supporting and enforcing co-ordination between social actors. Thevenot’s suggestion is that these arrangements are both cognitive and normative, that is to say, they require both sense-making and rule-making. Subsequent work with Boltanski, clarified this dual role of standards:
When persons grasp events as human actions in the perspective of coordination, they relate behaviours to some relevant good, the format of the good being highly variable ... When properly formatted, persons and things qualify for a certain mode of coordination. (Thévenot, 2006: 111-2)

Re-articulating the traditions of conflict and functionalist sociologies, Boltanski and Thévenot (2006) argued that in situations of conflict and uncertainty, actors and groups enter a dynamic of critique and justification searching to build the conditions for the qualification of people or things in orders of worth, which recursively act as cognitive scaffolds to understand situated action and act or follow rules appropriately. Importantly, they argue that actors’ moral justifications entail not only processes of sense-making but also, and importantly, require investment – durable, explicit, objectified commitment – to one way of ‘grasping events’.

Although not fully adopting Boltanski and Thevenot’s emplacement of the dynamics of critique and justification within a set of ‘grammars of the common good’ (cities), we draw on their model to support our empirical description of controversies about ways of measuring health From this perspective, standards such as DALYs are not mere boring, ‘technical objects’; instead, studying their making provides direct insight into institutions building, and the embedding of normative and epistemic arrangements. Our argument is further that there are not made de novo, but encounter a world that is already formatted. The conventional sociological framework to understand these processes is to see them as adjustment between ‘universal’ standards and ‘local’ practices, i.e. between macro and micro scales. Boltanski and Thevenot’s pragmatic sociology challenges this framework by drawing analytical attention to critical situations – controversies- as settings where the relationship between structure and agency is articulated and negotiated by social actors themselves, by striving for equivalence.

This is consequential for our approach to health metrics. Firstly, we must attend to a principle of symmetry (Bloor, 1991) to avoid falling into what Boltanski and Thevenot would call a ‘critical sociology’ of global health. This is done not rise above political disputes but to trace and understand how existing institutional formats became entangled with particular ways of knowing (e.g. Shaping and Sheffer, 1985). Thus, in the paper, we draw on scientific literature and published reports on health measurement and metrics, to identify the network of researchers, institutions and policies associated with different approaches to measuring
population health. We ask: how and why did particular forms of measurement become entangled with specific health care policies? In the second section of the paper, we link existing critiques of the normative assumptions embedded in the DALY to an important and outstanding controversy in the field of public health (Sassi, 2006; Voigt and King, 2014). Dating back to the turn of the 1980s, the controversy opposed what came to be known as the ‘primary health care movement’ to the proponents of an approach to international health care programmes that favoured ‘rationing’. We suggest that while the former relied on an evolutionary modelling of the relationship between types of public health programmes and the health of populations, the latter flattened the historical differences between societies to map disease and prioritise ‘interventions’.

As a result, we provide an analysis of health metrics that is saturated with diversity. Instead of seeing the consolidation of the DALY as a process unique to late modernity’s financialisation of health, in the third section of the paper, we reconstitute the process through which the DALY emerged within an older, historically constituted ecology of health measurement. We thus suggest that the DALY can be emplaced within what Clarke (1991) would conceptualised as a ‘social arena’ of health measurement: a field of action around a specific issue shaped by conflict, competition, cooperation and exchange. This provides insight into the relational underpinnings of the DALY’s genesis and development. By opening the analysis of health metrics in global health to this wider field of action, the DALY can be understood as the contingent outcome of an on-going process of negotiation. By exploring how ‘it could have been otherwise’ (Bijker and Law 1992), our paper opens the imagination to path and possibilities of alternative ‘investment in form’ in health.

1. Measuring Global Health

Historical analyses usually agree that the transition from international to global health is evidenced in the consolidation of a multi-actor approach to financing, organising and delivering health technologies and interventions in the 1990s, often underpinned by a collective belief that health can stimulate economic development (Packard 2016). For many, this transition was marked in two key publications: the 1993 World Bank World Development Report: Investing in Health and the 1996 Global Burden of Disease (GBD) Study report (Murray and Lopez 1996). In this, the World Bank report is seen to have set the reform agenda for healthcare system in the developing world, arguing for a ‘modernisation’ strategy underpinned by the overall goal of efficiency delivery through “greater diversity and competition” (World Bank 1993, iii). Seeing efficiency as driven by the optimal distribution
of resources obtained from competition between alternatives, the World Bank further advocated that this aim could be obtained by maximising the ratio between input resources – at a specific cost - and a specific health outcome (World Bank 1993, 25). The GBD Study, which the World Bank funded along with the World Health Organisation, proposed that to attain this efficiency goal it was necessary to develop a ‘single measure’, an “internationally standardized form”, that would support “policy-makers to compare the relative cost-effectiveness of different interventions” (Murray and Lopez 1996, 6-7), as well as compare health care systems across space and time. In focusing on developing a metric, they were motivated by the view that countries’ main burden of disease could no longer be straightforwardly divided along a communicable/non-communicable disease line, and that the “so called ‘epidemiological transition’ [was] already much further advanced [in the South] than many public health specialists appreciate[d]” (Murray and Lopez 1996, 1).

In 1971, the epidemiologist Abdel Omran proposed that the processes of demographic transition – from populations with high fertility and mortality towards low levels of fertility and mortality -, had been paralleled by a change in the dominant cause of death in the population, it being first characterised by ‘famine and pestilence’, then by ‘pandemics’, and finally by ‘degenerative diseases’ (Omran 1971). Although a model mainly designed to argue for the importance of public health infrastructure in producing the shifts in mortality as conceived by demographic transition theory (see below), the paper quickly became a reference point for emerging controversies and debates about the nature of the relationship between mortality and morbidity – the quantity of quality of life - in high-income countries (Weisz and Olszynko-Gryn 2009; Moreira 2016).

Such emphasis on the power of metrics in propelling the transformation of health care across territories and populations has recently attracted the interests of scholars from anthropology, sociology and science studies. Focusing on the role of such metrics in shaping of the contemporary governance of global health, these studies enable, in turn, the identification of the normative assumptions, local implications and institutional consequences of standardised measures of global health.

Drawing on a Foucauldian framework, Wahlberg and Rose (2015) explore the GBD Studies as a new technology of the governmentalization of life through the quantification of disability. Noting a shift in the forms of knowledge produced about populations, from mapping disease patterns to calculating the state of world health, they argue that this denotes
a new epidemiological style of thought in which “the problem of morbid death gives way to that of morbid living, made calculable through a metrics of ‘severity’, ‘disability’ and ‘impairment’” (Wahlberg and Rose 2015, 60). In this process, the social and personal consequences of living with disease become an object of political concern that is made knowable, calculable, and thus becomes a target of intervention. This new style of thought, they argue, drawing on GBD documentation, can be traced back to the development of the concept of ‘epidemiological transition’.

In 1971 Daniel Sullivan had published what is considered a seminal contribution to the measurement of health by proposing an index of ‘health expectancy’, a metric that calculated ‘Disability-free Life Expectancy’ (Sullivan 1971). Designed as a ‘rational and empirical formula’ to assist decision-making when “public health programmes compete with each other and with other government activities for budget allocations” (Sullivan 1966: 1), the Sullivan method is a key exemplar in a more general shift towards the measurement of quality of life from the 1970s onwards (Bowling 2005). Addressing this shift, Armstrong and Caldwell (2004) argue that the rise of this ‘quality of life’ repertoire can be seen as counterpoint to “the dream of a technological future” in health care. By measuring function or subjective well-being, such measures bring to bear the often modest benefits of biomedical technologies in the lives of patients. This being said, it is important to recognise that this questioning was closely linked to an observed stabilisation of mortality levels in the populations of high income countries in the 1960 and 1970s (Manton 1992), and the shared belief that only technological advances in chronic disease management – or ‘degenerative diseases’ – would enhance survival rates (Moryama 1964).

For Wahlberg and Rose, these new health metrics are equally encased by what they label a ‘new epidemiological gaze’ where the focus moves from “biological processes and events of life (disease and death) to social processes and events of living (disability and health)” (Wahlberg and Rose 2015, 86) in the management of populations. They thus see continuity between the health measurement movement of the 1970s, briefly described above, and the later work developed by the World Bank and the GBD Studies. While this might be justified by the fact that the key metric proposed by the GBD studies – the DALY – can be seen as an adaptation of the Sullivan method, there are significant technical differences between the metrics used. Usually, Sullivan health expectancies are measured by computing different states of health to obtain an aggregate of ‘full health’. By contrast, the DALY focuses on measuring the effects of a disease in reducing “human function”, one DALY being equivalent
to a ‘one lost year of full health’. Epistemically, the DALY can be grouped with measures of mortality and morbidity that focus on Years of Life Lost (YLL).

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Formula/definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>YLL</strong></td>
<td>YLL = N x L where: N = number of deaths L = standard life expectancy at age of death in years</td>
</tr>
<tr>
<td><strong>YLD</strong></td>
<td>YLD = I x DW x L where: I = number of incident cases DW = disability weight L = average duration of the case until remission or death (years)</td>
</tr>
<tr>
<td><strong>DALY</strong></td>
<td>DALY = YLL + YLD. The disability-adjusted life year (DALY) is a measure of overall disease burden, expressed as the number of years lost due to ill-health, disability or early death.</td>
</tr>
<tr>
<td><strong>QALY</strong></td>
<td>QALY = Years of Life x Utility Value The years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality of life score.</td>
</tr>
</tbody>
</table>

Table 1. Burden of disease metrics

Introduced by Mary Dempsey in 1947, with the aim of comparing mortality due to tuberculosis with heart diseases and cancer (Romeder and McWhinnie 1977), YLL specifically focused on causes of mortality for setting priorities in health care – those causing more mortality being higher in the ranking of concern and investment. Regularly published by national statistical agencies by the 1970s, YLLs were relatively stable set of procedures in epidemiology, not only because it used the central epistemic referent in epidemiology – death – but also because it could rely on readily available data on causes of death. It is from this perspective that it can be said that the DALY “extends the notion of mortality gaps to account
for time lived in health states worse than ideal health” (Murray, Salomon and Mathers 2000, 983).

Despite these conceptual differences, Wahlberg and Rose propose that there are key commonalities across metrics in that they enable forms of power that are conceived and implemented as “strategies for prioritization and evaluation” (Wahlberg and Rose 2015, 86). This is significant because it links the quantification of health to a mode of public health intervention that enacts healthcare as specific technological packages and programmes, whether to maximise health expectancy or to minimise health gaps.

From the same governmentality approach, Kenny (2015) analyses the design of DALY and its role in the governance of global health. Focusing her analysis on the World Bank’s formulation of the DALY, Kenny argues health metrics accomplish an ‘economization of life’ by disaggregating lifetimes into component units of time. Through this procedure, it is possible to represent life as a revenue stream to be maximized through practices of self-investment in one’s own health. In this, health is configured in terms of human capital expressed in the use of age weighting, discounting rates, disability severity weighting and standard life expectancy, the key procedures used in the DALY approach to derive measures of ‘lost health’ (Kenny, 2015: 12).

This linkage to the concept of ‘human capital’ provides Kenny with a bridge between health metrics and the wider transformations of what Foucault (2009) labelled neo-liberal forms of knowledge-power, and in particular by how the theory of human capital enacts labour as a combination of ability and investment (education, etc.) from which an income can be obtained. Human capital, as understood in economics, is the stock of skills that people possesses. This stock increases the value of people as labour force (Becker, 1962:9). “The flow of these skills is forthcoming when the return to investment exceeds the cost (both direct and indirect)” (Goldin, 2014: 53). Human capital encompasses the notion that investments in people such as education, training, health increase an individual’s productivity and therefore value. From this perspective, Kenny proposes that health metrics can be understood as ‘technologies of governance’, by reducing uncertainty on the types of investment – health technologies, or services – that produce health and maximise the value of individual labour (Kenny, 2015: 12).

The consequences of framing health in such a narrow way is brought to bear in Adams’ (2016) ethnography of the impact of audit practices and quantification in global health.
Emphasising the role of international organisations, NGOs and governments in this process, she argues that they have led to a dual economization of health and fiscalization of life. Similarly to Kenny, Adams sees the economization of health as resulting from the adoption of ways of organizing and thinking about healthcare from economics, particularly those based on understanding healthcare as a market. The fiscalisation of life is related with the increasing participation of actors from the financial sector in global health. In these processes, health metrics are key instruments in the economization of health. Furthermore, different health metrics also enable the maintenance of distinctions between economic regions of the globe, the emphasis on ‘quality of life’ (QALY) being distinguished from that on ‘disability’. Such distinctions make sense, Adams argues, because “in the Global North what mattered was counting the cost of keeping people alive, whereas in the Global South what mattered was plausible justification for continued expenditures in relation to death and disease burden” (Adams 2016, 27).

Adams major concern is the side effects of these calculations. By emplacing healthcare in a market framing, countries, communities and healthcare providers have to adjust to the demands of the production of data. Thus, the production of data about impact becomes the major goal of strategies and funded projects rather than the actual delivery of healthcare. This exerts violence, particularly through the necessary arithmetic gymnastic to make countries comparable, and the pretention of value neutrality in resource allocation that these metrics make possible (Adams 2016, 30). Finally, for Adams, the most serious consequence of the economization of health is to think of life as money, substituting politics for a market principle. This renders marginal and invisible other forms of evidence and other experiences in global health that –despite working in practice–cannot be translated into the language of market (Maldonado, 2018). This ‘silencing’ of local ways of enacting and valuing health practices endangers the sustainability of global health programmes and undermines health equity as a principle.

Drawing on Foucault’s writings on governmentality, Wahlberg and Rose, Kenny and Adams provide an important set of insights on global health as an apparatus of knowledge–power, highlighting the historically constituted, political and economic contexts in which it emerged. Their analysis is particularly relevant to understand how specific modes of knowledge making and evaluation of evidence became uniquely suited to support forms of governance that seek to maximise efficiency in the delivery of discrete interventions and technologies. This brings to bear the contingent nature of current global health processes and institutions,
while providing important clues on their genealogy. However, a disadvantage of this analytical approach is that, by departing from the central role metrics such as the DALY play in contemporary global health, it tends to understate the controversies and indecisions that led to the present situation. To recover the paths not taken - to fully account for ways of enacting and valuing health that have been silenced – it is necessary to focus on the institutional transformations of the governing of international health during the second half of the twentieth century and ask the question: What other forms of global health policy could have been possible?

2. Recovering diversity in global health: social change vs. efficiency

In this section, we aim to recover some of the epistemic and political diversity that preceded the stabilisation of the global health apparatus as we know it today. To do this, we shift our focus from the establishment of this assemblage in the 1990s to exploring a wider, longer process. As suggested by Cueto (2004), it is possible to link the focus on metrics in global health to a controversy about the role of primary health care in international health policy in the late 1970s (See also Brown et al. 2006; Rodin 2013).

As we argued above, in the 1970s, experts such as Omran aligned with policy makers and the World Health Organisation (WHO) to propose a focus of international health on primary health care (Cueto, 2004). Through the motto ‘Health for All’, the primary healthcare movement agenda expressed a social understanding of health, and a belief that policy should be focused on the improvement of living and material conditions of the poor. It proposed to deploy community-embedded physicians as agents of change. Support for this position came both from acceptance of social epidemiological models of public health such as McKeown’s (1976) and from “a new political context characterized by the emergence of decolonized African nations and the spread of national, anti-imperialist, and leftist movements in many less-developed nations” (Cueto 2004, 10). This political orientation was expressed in the Alma-Ata declaration of 1978 which identified as the principles of the primary health care movement the use of appropriate technology, the opposition to medical elitism, and a concept of health as a tool for socioeconomic development (Cueto 2004, 12). A year later, the 32nd World Health Assembly endorsed the conference’s declaration.

In the context of the Cold War, however, the Alma-Ata declaration was not well received by the main international donors, particularly by the United States. The declaration was accused of being abstract and too holistic to define specific, grounded policies (Schwartländer, 1997).
Key amongst the critics was the Rockefeller Foundation’s concern with the lack of strategies to define priorities in international health. This led to it sponsoring a small conference in 1979 entitled “Health and Population in Development” at its Bellagio Conference Centre in Italy. The conference gathered representatives from the Work Bank, the Canadian International Development and Research Centre; the Ford Foundation and the US Agency for International Development (USAID). As its main outcome, the conference led to the production of a report written by Julia Walsh and Kenneth S. Warren (1979) where an alternative set of principles to the Alma-Ata conference and the World health Assembly were outlined.

Entitled Selective Primary Health Care, an Interim Strategy for Disease Control in Developing Countries, the Bellagio Centre conference report is key to understanding the stabilisation of the global health assemblage as we know it today (Walsh and Warren, 1979). Casting the objectives of the Alma-Ata conference as “laudable but unattainable in terms of its prohibitive costs and the numbers of trained personnel required” (Walsh and Warren 1979, 145), the authors proposed that the concept of selective primary health care should be underpinned by the need to set priorities in what health programmes and interventions to finance. This entailed two sets of data generation: identification of target diseases, based on assessments of prevalence, morbidity, and risk of mortality; and evaluation of the medical interventions used to manage those illnesses, which was to be supported by measures of effectiveness and cost. Primary healthcare, the report argued, should focus on cost-effective interventions, those that reduce most of the burden of a particular disease – here measures by numbers of deaths - to the lowest cost. This approach was thus to be infrastructurally supported by a combination of measures of differential contribution of specific diseases to the overall mortality, with assessment of effectiveness produced through randomised controlled trials of specific interventions.

Their proposals came to bear one year later in the Kaiser Foundation International-sponsored Ghana Health Assessment Project (GHAPT 1981). Seen as a key technical antecedent to the formulation of the DALY and the GBD study of 1993 (Murray, Salomon and Mathers 2002), the project was explicitly motivated by the apparent lack of progress in levels of population health despite increases in health care facilities and personnel in Ghana. To investigate this problem, the investigators designed a method to estimate the health impacts of different diseases, thus enabling the examination “of the magnitude of the change [health interventions] might be expected to produce in health status” (GHAPT 1981, 73).
Recognising the technical difficulties of measuring the impact of illness in economic activity, the Project team opted for a measure of ‘healthy days of life lost’ as a proxy assessment of the impact of illness on productivity. This form of measurement, in turn, was deemed to enable the formulation of priorities by quantifying the “healthy life saved” by, for example, a measles immunisation program vis-à-vis an outpatient clinic for the same condition. These arguments, it was suggested, should be more “persuasive...than the mere assertion that prevention is better than cure” (GHAPT 1981, 77).

This contrast between ‘mere assertion’ and ‘quantification of healthy life’ makes reference to the controversial nature of the Ghana project team and the Bellagio Centre conference’s proposals. Although not explicitly challenging the alternative position, it aimed to grade the epistemic quality of different approaches to prioritisation. In so doing, it provides, albeit from a biased standpoint, the link between styles of reasoning and knowledge making about population health and the politics of health care that underpins this controversy. The conflict was thus about the politics of knowledge in international health.

On the one side, there was the loose alignment of institutions, agencies and actors formalised the Alma-Ata declaration, a coalition that until the mid-1970s included the World Bank itself, in its support for prioritising of sanitary conditions and housing “as causes of poor health” (World Bank, 1975). They proposed that priorities should result from an understanding of basic rights and key principles. Indeed, the Alma-Ata declaration recommended that to be able to achieve a socially and economic productive life it was a necessary condition that citizens were given access to basic sanitation and water supply, immunization against major infectious diseases, maternal and child health, health education, etc. (WHO 1978, 2). Contrary to the Ghana’s team claim, this was not a ‘mere assertion’: it was based on established ways of analysing mortality and morbidity data, perhaps best epitomised in Omran’s model of epidemiological transition. Omran’s model drew on the developmental narrative of demographic transition theory, which had underpinned many of the fertility programmes implemented in the ‘third world’ by the UN, to propose that “mortality is a fundamental factor in population dynamics” (Omran 1971, 511). The model suggested that while transition to the ‘age of receding pandemics’ had been sustained by the introduction of basic sanitation and other infrastructural changes, the shift to the low level equilibrium of the ‘age of degenerative diseases’ could only be achieved by widening access to basic health care provision (maternal health, etc) (Weisz and Olszynko-Gryn 2009). This developmental way of reasoning about populations, if not directly the theory itself, provided, for those supporting
the Alma-Ata declaration, a sound, robust basis for the elaboration of priorities and public health programmes that fitted the needs of specific populations.

On the other side were those that, as described above, argued that the scarcity of resources available dictated the need for a different way of setting priorities in international health. By the late 1970s, however, these were in a relatively marginal position (see WHO 1978). To overcome this, proponents of the selective approach had to challenge two aspects of the Alma-Ata line of reasoning. Where the latter saw health as a driver of economic wealth creation, the Rockefeller Foundation and the World Bank position was that available economic wealth constituted an antecedent restriction to investment in health (McNamara 1978). By highlighting the shortage of resources limiting the formulation of policies, they were returning to Malthusian-type arguments about the regulatory role of scarcity in the dynamics of populations, which demographic transition models had originally attempted to displace (Ramsden, 2002). Now encased in the language of mainstream economics, scarcity entailed the formulation of choices, with differing opportunity costs. But how were those choices to be articulated? Answering this required a critique of the epistemic basis used by those supporting the Alma-Ata declaration. As Walsh and Warren put in their report on the Bellagio Centre conference:

_The state of ill health found in many less developed countries should not be approached as a single problem. Traditional indicators such as child mortality or life expectancy are insufficient for grasping the issues involved. Health is complex multifaceted problem, an amalgam of many diseases with multiple causes. Indicators are actually distilled composites of hundreds of different health problems and disorders. [...] Diseases endemic to the less developed countries are protean in their etiologies, mechanisms of transmission, impacts on humans and susceptibility to attack. It is unlikely that any single mode of would be suitable for all. Each disease must be considered individually, with its unique mix of epidemiological, ecological and social factors. (Walsh and Warren 1979, 145-6)_

Instead of patterns of mortality related to general types of disease (e.g. infectious diseases), Walsh and Warren proposed that morbidity could be disaggregated into different ‘etiologies’. This made ‘indicators such as child mortality or life expectancy’ inadequate to capture not only the mechanism of disease causation but also the impact of ill health on persons. They
were thus also ill equipped to serve policy formulation. Such a disparaging assessment of the value of life expectancy to understand population health is of key significance. Although at least since the 1960s, as discussed above, researchers had been trying to develop metrics that could capture the ‘quality’ of life lived, most notably by Sullivan (1970), these, as the GHAPT put it, “did not relate […] to specific diseases and thus [were] not likely to be of direct value in determining the allocation of health resources” (GHAPT1981, 77). The indicator that most suited such purpose, they argued, was instead the concepts of ‘healthy life lost’ to illness, which had originally been developed by Dempsey in 1947 to question the value of mortality rates in creating the emergent view that tuberculosis was decreasing its significance in the health of the US population (Romeder and McWhinnie 1977).

Further, it was an accepted view amongst those attempting to develop measures of population health that calculations of ‘quality of life’ were best suited for societies where the prevalence of chronic disease justified a focus on the impact of illness on social functioning (e.g. Sullivan 1966, 1). These were assumed to map onto societies in the ‘developed world’. But, during the first half of the 1980s, the WHO Global Epidemiological Surveillance and Health Situation Assessment Unit led a series of studies that questioned the neat mapping between development regions of the world and the prevalence of types of disease. For example, using WHO mortality data, it was found that “50% of all deaths in the world due to [cardiovascular illness or cancer] occur in the developing world” (Hakulinen, Hansluwka, Lopez, and Nakada 1986, 347). Using mortality data was important because it directly challenged the epidemiological transition model with its own preferred statistic. This challenge did not lead to a rejection of the epidemiological transition theory, however. What appears to have happened is that, for those supporting a selective approach to international health care programmes, the model came to refer to overlapping patterns of mortality observed in developing countries, losing its explanatory power along the way (e.g. WHO 1991, 75). The shift was such that by 1994, the GBD lead researchers could propose that, in developing countries the “so-called ‘epidemiological transition’ is already much further advanced than many public health specialists appreciate” (Murray and Lopez 1994, 2: our emphasis).

To a very significant degree, this controversy is still unresolved (Shreker 2015). One the one side, there are those, endorsed by the WHO and supported by the Melinda and Bill Gates Foundation, who argue for the need to validate new methodologies and implements to ascertain the value of health interventions across populations. This can be seen as epitomised in the establishment of the Institute of Health Metrics and Valuation and the GBD 2010
study, which aimed, as referred above, to measure the quantity of health in terms of the ‘limitations’ illnesses imposes “on what [persons] can do in life” (Salomon et al. 2013). They favour methodologies that aim to identify ‘gains’ in health obtained from the application of discrete interventions, and simulate the ideal of economic choice of investment in the face of scarce resources. On the other side, there are those who argue that existing statistical techniques enable the identification of constant conjunctions between socio-economic conditions and public health infrastructure, and health outcomes across populations (Scribner et al. 2017). They thus place less value on developing a standardised measure of health gain, and more on making visible and quantifying the inequalities and inequities in global health. For them, health research should not be concerned with the question of whether public health interventions work but with understanding the process of making them work in specific contexts; if is not a question of if but how.

3. Context means comparison
In the first section of the paper, we have suggested that contemporary social science research has identified a key role for health metrics in the governance of global health processes. In this, there is some agreement that health metrics support understanding and managing health care through an economic framing, both in terms of how health is seen to lead to economic development and how decision making is deployed as a discrete choice between alternative ‘goods’. We have argued however that this analysis partially glosses over the uncertain and controversial nature of using DALYs to allocate health care resources across populations. Such a proposal is better understood against an agonistic background, where divergent ways of linking knowledge making on population health to health care policy and institutions have maintained a dynamic of critique and justification for at least three decades.

The key consequence of this analysis is the realisation that health metrics should not be investigated in isolation. As was suggested on both sections, DALYs belong to a wider set of health measurement techniques and procedures; while the DALY has been the main form used to measure health in the Global South, its generation and transformation is integrated in a wider history of health measurement and valuation, as was indicated above. Health measurement assembles what Clarke (1991) would label a ‘social arena’, The importance of this wider field of action was, in fact, acknowledged by the developers of the DALY themselves in the GBD publication that first described the metric:
In order to capture the impact of both premature death and disability in a single measure, a common currency is required. Since the late 1940s, researchers have generally agreed that time is an appropriate currency: time (in years) lost through premature death, and time (in years) lived with a disability. A range of such time-based measures has been developed in different countries, many of them variants of the so-called Quality Adjusted Life Year or QALY. For the GBD, an internationally standardized form of the QALY has been developed, called the Disability Adjusted Life Year (DALY). (Murray and Lopez 1996, 6-7)

Opposing the use of a common currency to the ‘partial and fragmented’ health statistics of specific causes of death, Murray and Lopez emplace their new metric within a longstanding field of research where there is consensus that time (years) should be deployed as basic unit of measurement. This consensus, they suggest, has enabled the generation of a variety of measures which can be grouped around the idea of Quality Adjusted Life Year, of which the DALY is a variant. Thus, it is proposed, it represents an ‘international standardised form of the QALY’.

Given what was discussed in relation to the DALY in the previous section, Murray and Lopez’ assertions might appear unusual. This is because it is now accepted that, if anything, the DALY represents the negative of QALY, measuring disability or ‘healthy life lost’ instead of quality of life, or as Murray, Salomon and Mather (2000) would put it years later measuring ‘health gaps’ rather than ‘health expectancy’. However, in the moment when Murray and Lopez were writing the QALY represented the most widely known metric used in health. As such it constituted a point of reference, a signifier for an ‘arena’ into which the DALY could be placed. This was important for GBD lead researchers because it enhanced the DALY’s differentiation from the ‘partial and fragmented’ health statistics used by public health researchers and advocates, i.e., those who relied on ‘epidemiologic transition’ arguments to justify policies, as argued in the previous section. Like the QALY, the GBD researchers argued, the DALY aimed “to disentangle epidemiology from advocacy in order to produce objective, independent […] assessments of” health (Murray and Lopez 1996, 6). Linking the DALY to the QALY enabled thus the establishment of a boundary between the previous work done by epidemiologists comparing the health of populations, such as Omran, and the work of the WHO GBD studies team. Crucially, like the QALY, it did such
assessments in order to “allow policy-makers to compare the relative cost-effectiveness of different interventions” (Murray and Lopez 1996, 6).

However, GBD researchers were attempting to differentiate between QALY valuations that depended not only on different techniques of quantification but also on how different societies valued health, and a form of valuation that enabled comparisons between territories and populations. This aim hinges on the ability to craft what Mol (2002) labelled a ‘comparative platform’, an operation supported by techniques that place – side by side – differences and similarities between cases (Moreira 2007; Stengers 2011; Deville, Guggenheim and Hrdličková 2016). In so doing, it creates that which is compared, that is to say, comparison elicits value that is commensurate. However, rather than producing commensuration by negotiating and compromise, as did QALY proposers, the DALY was firmly attached to a set of ‘value judgements’, which Murray and colleagues deemed to be ‘equalitarian’. Such value judgements underpin the technical basis of valuation in the DALY approach: its focus on time lost, the emphasis on healthy life lost in the proximate future, the use of standardised disability weights and the differential valuation of time lived at different ages. The latter two are particularly illustrative of the process of standardisation underpinning the DALY.

Regarding age-weighting, instead of linking differential valuation of disability by age to ‘productivity weights’ as would be predicted by the theory of human capital, DALY champions viewed “unequal age-weights as an attempt to capture different social roles at different ages” (Murray 1994, 435). Moving away from a strictly economic valuation of social functioning, suggested for example by the World Bank (1993) (see above), they emplaced value judgement on the ‘original position’ – the ‘thick veil of ignorance’ – a thought experiment proposed by Rawls (1971), and much used in the health policy thinking (e.g. Menzel 1999). Interestingly, to operationalise this experiment, instead of deriving weights by application of general principles of the relationship between age and role, they conducted a Delphi conference with ‘public health experts’ so as to obtain a function of weighting by age.

In relation to disability weighting, although using the similar consensus generating techniques, the procedure hinged on a specific normative differentiation that would prove problematic. The WHO International classification of impairments, disabilities, and handicaps (1980) differentiated between ‘disability’ – a functional limitation resulting from
an organic impairment – and ‘handicap’ – a socio-economic disadvantage or impact resulting from functional limitation. Recognising the attraction linked to measuring handicap, DALY proponents were concerned that it might lead to reinforcing inequalities as the ‘same’ functional limitation would be valued differently depending on the context where people lived (Murray 1994, 437-8). This decision had the key advantage of enabling weights to be generated without knowledge of the particular circumstances experienced by individuals, disability being defined by ‘universal’ human abilities such as seeing, speaking, grasping and activities such as dressing or washing.

Both age-weighting and disability weighting proved to be, for different reasons, especially problematic features of the DALY (Voigt and King 2012). Age weighting was attacked for relying on the assumption that the disabling effects of disease become less significant with age. It was not only that procedure made widespread suppositions about the value of older people in society across cultures, but also that it was going against the emerging scientific and political questioning of chronological age as valid marked to index individuals function or health (Moreira 2016). As in the domain of disability politics, activists argued that such beliefs effectively prevented older people from participating in social and economic life.

Indeed, by the turn of century, the World health Organisation itself had articulated a policy framework – ‘active ageing’ – that explicitly aimed optimise economic, social and cultural participation throughout the life course by removing barriers to participation, including in inclusion of older people in research (Lassen and Moreira 2014). Response to this shifting context, and to challenges from bioethicists (see Murray, Salomon, Mathers and Lopez 2002, Part 13), entailed making age weights progressively uniform in the GBD 2010 (Voigt and King 2014, 226).

However, criticisms related to disability weights proved more challenging as it questioned the very possibility of comparing levels of health across populations and time points. While the solution to the age-weighting problem reinforced the internal consistency of the standard, disability politics threatened to undermine the key comparative aims of the DALY (Murray and Lopez 1994). To address this problem, new weights were defined through surveys in which lay descriptions of different health states were used to elicit valuations (Ustun et al 2001). These were not seen as reflecting preferences, as in the QALY approach, but instead by appealing to an ‘intuitive, common sense’ notion of health (Murray, Salomon, Mathers and Lopez 2002, 736-7) which allowed cardinal grading of health states. It was then possible
to map differentiated weights onto territorially bound prevalence of disease and disability data. But the aim remained to isolated loss in health from disadvantages in welfare and from the impact of the social environment. In this, the goal was to measure ‘health’ in a narrow sense, independently of consideration about welfare and socio-cultural differences.

Such belief in the ability to isolate health from welfare and social context is precisely the major criticism make Voigt and King to the GBD 2010. They argue that not only is the belief ill founded, as health is interdependent on social, economic and cultural conditions (Blaxter 2004), but also that it entails a narrowing of the criteria for political making decision making. They suggest that while using metrics such as the DALY might be key to monitoring health equity across the globe, this should be distinguished from “using them to determine priorities and allocate resources among different diseases in different countries” (Voigt and King 2014, 227). However, as argued above, the comparative epidemiological aim cannot be easily separated from the politics of decision making. For GBD researchers, mapping of disease burden and valuation of ‘healthy life lost’ is aimed at shaping, or disciplining how decisions are made. This, in effect, requires foregoing complexity for commensurability or comparability. As the GBD lead researchers put it:

Comparability is fundamental to the use of survey results for development of evidence for health policy but has been under-emphasised in instrument development [...] The basis of science is comparable measurement: comparison creates possibilities of investigating broad determinants at national and cross-national level. [...] Health measurements, particularly for policy makers, generally only have meaning in context, and context means comparison. (Murray, Salomon, Mathers and Lopez 2002, 746-7)

Context means comparison. By presenting comparability as the condition for good, rational and effective decision-making, the GBP researchers situate local and national policy in a scenario in which the international context guarantees objectivity. Context is, in this case, the type of information available for the decision maker. As with the QALY, health measurement is defined in relation to an imagined type of decision making for which normative assumptions are made. In practice, however, such assumptions operate through infrastructures in which measurement and evaluation are the gateway for funding and programmes justification.
4. Conclusion

Health metrics can be conceptualised as a social arena (Clarke, 1991). They have been shaped by reference to each other and by common conflicts, connections and exchanges. These metrics have shaped Global health as a domain of research, policy and practice. A closer view to the development of metrics such as DALY has shown the interactive relation between health measurement and politics of health. This relation has been characterised by uncertainty and debates between epidemiologists and other global health experts about the consequences of measuring in health decision-making.

This paper has presented some elements to understand the entanglement between metrics and specific health policies. Measurement and quantification in health has been a contested field in which are debated, negotiated and legitimated conceptions about health and social policy. Our starting point was the social science scholarship about global health metrics. These works have shown how the market as normative model has shaped health policy and the key role of metrics in this process. However, as we have presented, this account is partial and does not take into consideration the contested development of these metrics and the debate in public and global health about their normative assumptions. We have presented two facets of this process: the roots of DALY in the controversy about Primary health care and the epistemological connections between health. The metrics are health currencies based on time.

The controversy about Primary health care has shown the political interest in establishing a monopoly around measuring and quantification in healthcare. Despite the Alma-Ata formulation was based on the epidemiological transition model, its critics claimed that the lack of the right measuring infrastructure made the declaration abstract and idealistic. DALY and the Global Burden of Disease are development of a critique that introduced scarcity in health measuring and decision making. Scarcity will reach a regulatory role, encased in the language of mainstream economics. The discussion about costs of opportunity and cost-effectiveness that scarcity entails are put in the centre of health measurement.

On the other hand, this paper has examined the origins of DALY as an international standardised version of QALY and the epistemology that these metrics share. The development of the DALY and the Global Burden of Disease shows the use of standardisation and quantification as response to what is perceived, from a technocratic perspective, as conventional politics. However, as we have presented, such standardisation is
generative and have made possible new debates about the normative assumption of health decision-making. In attempting to address critiques the DALY has revealed its conventional basis and its limitations to capture the ‘value of health’ for specific populations.

The most remarkable feature of DALY is its comparability, across types of illness but most importantly across nations. The importance of comparison is a consequence of the interest of GBD researchers in establishing a concept of health independent of references to socio-cultural specificities and material context. DALY and GBD infrastructures are shaped by a commitment to capturing health as a ‘pure’ quality. Through comparability DALY provides its own context and becomes the basis of objective health decision-making.

We are witnessing the development of an institutional complex of metrics, policies and funding in Global health based on the idea that measuring ‘health’, in a narrow sense, independently of considerations about welfare and socio-cultural differences constitutes the best way of tracking the effectiveness of interventions and policies. These metrics provide a context for decision making, significantly affecting how policies, interventions and resources are justified and contested. Our paper has hopefully contributed to open the framing of these decisions. By exploring how path not taken, our paper offers new possibilities for alternative standard and ‘investment in form’ in global health.

References


25


