Excerpt from The Tango of Ethics: Intuition, Rationality and the Prevention of Suffering (Imprint Academic, 2023) by Jonathan Leighton

14. From Ethics to Action 187

Creating a new suffering
metric for health economics

Returning to the very concrete and tangible: in order to determine priorities and measure progress, we need metrics that are based on what matters. I have argued that everything that matters ethically, in the sense of there being inherent urgency to change it, can be reduced to suffering. There are also strong intuitions, such as for the continuity of life and against knowingly causing harm, that, if not respected, create suffering as well.

The two most widely used health metrics, the QALY (Quality-Adjusted Life-Year) and the DALY (Disability-Adjusted Life-Year), are strongly based on the intuition towards preserving life. The QALY considers a full year of healthy life to be a reference, and a year of life lived at a lower level of health to be less valuable. The DALY measures years of life lost to death or "disability" (actually meaning departure from optimal health), and similarly considers a year of life at a lower level of health comparable to a year partially lost altogether. Health economists literally weigh the value of lives based on their quality adjustment and number of expected years remaining. QALY-based cause prioritisation has also been advocated for within the effective altruism movement (MacAskill, 2015).

When QALYs are used to weigh different outcomes in terms of QALYs lost or gained, they directly imply that the death of an older person is less costly and therefore more acceptable than the loss of a younger person, and also that the death of someone with a chronic illness is less costly and therefore more acceptable than the death of someone in good health. Although this might appear shocking—at least from a perspective that

considers all lives to matter equally, regardless of how healthy people are—this is exactly the way much of mainstream health economics works, and the acceptability of a policy is often argued to be a direct consequence of its impact on human lives, expressed in terms of QALYs. This means that, all things being equal, a system that relies mainly on this metric to make decisions will prefer to save the life of a 40-year-old over a 60-year-old, and the life of a healthy athlete over someone of the same age with diabetes. A newborn who has barely seen the light of day beats everyone else (with the possible theoretical exception—if the reasoning were to be extended—of a 3-month-old foetus, with a full lifetime of QALYs ahead plus a half-year of in utero hedonic zero).

Furthermore—and this may be more of a theoretical criticism than one based on common usage (Dolan, 2001)—if applied aggregatively to a population rather just to measuring outcomes for individuals, the QALY paradigm would have the absurd consequence of allowing severe suffering to be compensated by having more individuals alive experiencing that suffering! If someone has a terrible chronic illness that causes them to suffer greatly, then keeping them alive is less important according to the QALY paradigm, but if we could bring into being several such people, that might still be considered better than having just one healthy, suffering-free person. QALY-based reasoning could, in principle, thereby lead to the Repugnant Conclusion, and even justify a worse scenario where everyone is suffering severely.

The QALY, which is currently used by the National Institute for Health and Care Excellence (NICE) in the UK, among other places, was intended to be a useful metric for making health and economic decisions. But it is closely aligned with the notion that our goal is to create "value" in the universe, and that a human life is valued by how much happiness it will produce. As I've argued repeatedly in this book, I don't think this is a rational approach to achieving a better world—or the most humanising approach, either. In fact, in one way the QALY paradigm seems conceptually even more flawed than the classical utilitarian weighing of happiness vs. suffering, in that it doesn't even explicitly acknowledge suffering as inherently bad, but rather, as instrumentally reducing the positive value of a life.

The DALY, which is used by the Global Burden of Disease project and the World Health Organisation (WHO), is philosophically somewhat distinct. While the use of QALYs inherently values years of life as something to maximise, DALYs focus on what is lost from a normal life due to poor health. Yet even DALYs don't focus explicitly on suffering as a core phenomenon to measure and address. Although most diseases with a high disability score are accompanied by serious suffering, this is not

always the case. Anencephaly (where part of the brain is missing) is an extreme case of a fatal condition that is entirely disabling but appears not to cause suffering, but people can psychologically adapt to other, less disabling conditions and thrive, especially if there isn't severe pain. More concerning, diseases or conditions that cause severe suffering but leave the patient relatively functional most of the time have a disability weight that hardly conveys the seriousness of the suffering. The theoretical maximum disability weight is 1.0, but average disability weights for the range of conditions tend to plateau at about 0.6 or 0.7 (though for some conditions the range extends to a maximum of 0.8 or even 0.9), meaning that even some of the most disabling conditions are still only considered to reduce health by two-thirds, on average. Yet, in part also because of the use of aggregation, even a disability weight nearing 1.0 might not cause a rare condition to stand out in the overall rankings as a target for urgent relief of suffering.

In some cases, patients commit suicide to escape their suffering. Yet a DALY measure that tied the suicide to the condition would, in principle, be more affected by the years of life lost due to the suicide than by the actual suffering incurred while the patient was alive.

What we see, then, is that the metrics being used don't reflect the full range of things that matter, and neglect the parameter that actually matters most. In addition, the methodology for determining disability weights or life quality involves asking healthy people to compare diseases or conditions that they don't have. This is a very poor way of assessing severe or extreme suffering—a situation that many cluster headache patients have experienced when trying to convey the severity of their condition to others.

Health metrics are a major factor in government priority-setting. There is a need for a new metric that better tracks the most ethically relevant parameter, suffering. The proposed metric Wellbeing-Adjusted Life-Year (alternatively referred to as WELBY, WELLBY or WALY) is a step in that direction, in that it more explicitly addresses subjective wellbeing as the key parameter (Frijters et al., 2019). According to one definition offered by wellbeing economist Paul Frijters (2020), it is a "one point change in life satisfaction for one person for one year when measured on a 0–10 scale". But it doesn't substitute for a metric that can capture the agony of intense and extreme suffering. It also shares the fundamental problems of aggregation and hedon-like metrics with QALYs and DALYs—that wellness or life quality points are summed up, with changes considered equivalent no matter where they are on a scale of wellbeing, and potentially no matter how these changes are distributed among a population.

A major report by the Lancet Commission on Palliative Care and Pain Relief (Knaul et al., 2018) proposed a metric called the SALY (Suffering-Adjusted Life-Year), modelled on the DALY but based on suffering rather than disability or health. Since years of life lost do not directly contribute to suffering, SALY could exclude the Years of Life Lost (YLL) component and be reduced to Years Lived with Suffering (YLS), with a weighting based on degree of suffering. However, because of the inherent problem with aggregation, there is no adequate way of combining mild, moderate and severe suffering into one metric.

It would therefore be advisable to create additional metrics of severe suffering to be used in parallel to a YLS measure. Years Lived with Severe Suffering (YLSS) could capture suffering at the level of approximately 7/10 and above. A separate metric called Days Lived with Extreme Suffering (DLES) could capture the most urgent suffering at the level of approximately 9/10 and above, and properly account for it even when experienced on short timescales. The goal would not be to effect any major shift in resources away from patients suffering moderately, but to ensure that the most severe suffering is given greater visibility and treated with the urgency it demands.

Suffering metrics could also be applied to non-human animals, notwithstanding the practical difficulties in assessing their degree of suffering (Savoie & Sarek, 2018; Welfare Footprint Project, n.d.). This would allow them to be included in a more global assessment of priorities, on the basis of the one parameter that matters most. Adapting a human health metric to non-human animals might be unprecedented, but a focus on suffering would do away with any objections based on how much we, as external observers, happen to subjectively value human and non-human lives.