

Drew Carter, Jason Gordon and Amber M. Watt

**Competing principles for allocating health care resources**

The Journal of Medicine and Philosophy, 2016; 41(5):558-583

© The Author 2016. Published by Oxford University Press, on behalf of the Journal of Medicine and Philosophy Inc. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

*This is a pre-copyedited, author-produced PDF of an article accepted for publication in The Journal of Medicine and Philosophy following peer review. The version of record Drew Carter, Jason Gordon and Amber M. Watt, Competing principles for allocating health care resources, The Journal of Medicine and Philosophy, 2016; 41(5):558-583 is available online at: <http://www.dx.doi.org/10.1093/jmp/jhw017>*

**PERMISSIONS**

[https://academic.oup.com/journals/pages/access\\_purchase/rights\\_and\\_permissions/self\\_archiving\\_policy](https://academic.oup.com/journals/pages/access_purchase/rights_and_permissions/self_archiving_policy)

**Accepted Manuscript**

*The accepted manuscript (AM) is the final draft author manuscript, as accepted for publication by a journal, including modifications based on referees' suggestions, before it has undergone copyediting, typesetting and proof correction. This is sometimes referred to as the post-print version.*

Immediately upon publication authors may:

- Immediately upload their AM to their own personal webpage (excluding commercial websites and repositories)
- Immediately upload their AM to their institutional or other non-commercial subject based repositories on the proviso that it is not made publicly available until after the specified embargo period

**After** the embargo period authors may:

Upload their AM to institutional repository or other non-commercial repositories and make it publicly available. Accepted Manuscripts may not be uploaded or shared on commercial websites or repositories, unless the website or repository has signed a licensing agreement with OUP permitting such uploading or sharing.

**Embargo periods - The Journal of Medicine and Philosophy - 24 months**

Embargo periods may vary between journals. For details of a journal's specific embargo period, please see the information for each individual title on our [Accepted Manuscript embargo](#) page.

When uploading an accepted manuscript to a repository, authors should include the following acknowledgment as well as a link to the version of record. This will connect the published version to the AM version in the repository and help ensure that the article is cited correctly.

*This is a pre-copyedited, author-produced version of an article accepted for publication in [insert journal title] following peer review. The version of record [insert complete citation information here] is available online at: xxxxxx [insert URL and DOI of the article on the OUP website].*

**27 August 2018**

## Competing principles for allocating health care resources

*We clarify options for conceptualising equity or what we refer to as justice in resource allocation. We do this by systematically differentiating, expounding then illustrating eight different substantive principles of justice. In doing this, we compare different meanings that can be attributed to 'need' and 'the capacity to benefit'. Our comparison is sharpened by two analytical tools. First, quantification helps to clarify the divergent consequences of allocations commended by competing principles. Second, a diagrammatic approach developed by economists Culyer and Wagstaff offers a visual and conceptual aid. Of the eight principles we illustrate, only two treat as relevant both a person's initial health state and a person's capacity to benefit per resource unit expended: (1) allocate resources so as to most closely equalise final health states, and (2) allocate resources so as to equally restore health states to population norms. These allocative principles ought to be preferred to the alternatives if one deems relevant both a person's initial health state and a person's capacity to benefit per resource unit expended. Finally, we examine possibilities for conceptualising benefits as relative to how badly off someone is, extending Parfit's thought on Prioritarianism (a prioritising of the worst off). Questions arise as to how much intervention effects accruing to the worse off count for more, and how this changes with improving health. We explicate some recent efforts to answer these questions, including in Dutch and British government circles. These efforts can be viewed as efforts to operationalise need as an allocative principle. Each effort seeks to maximise in the aggregate quanta of effect that are differentially valued in favour of the worst off. In this respect, each effort constitutes one type of Prioritarianism, which Parfit failed to differentiate from other types.*

*Keywords: health care rationing, resource allocation, health policy, principle-based ethics – social justice*

## I. INTRODUCTION

Aspirations to allocate resources, especially health care resources, ‘according to need’ have long been advanced. Equally, however, these aspirations have been subject to ambiguities concerning the meaning of need. Ambiguities have been more effectively addressed in the case of allocation ‘according to expected benefit,’ especially by health economists, with considerable consequences for health policy. Recent articles (Hope, Osterdal, and Hasman 2010; Carter et al. 2013) have attempted to advance thinking on the meaning of need when used as an allocative principle. In this article, we continue that effort, as part of a more comprehensive analysis. We systematically differentiate, expound and illustrate eight allocative principles that can be viewed as competing for our reflective endorsement and practical application. One of the aims of this analysis is to differentiate and set out, in quite general terms, different meanings that can be attributed to ‘need’ and ‘expected benefit’ (or ‘the capacity to benefit’). Another aim is to draw together works in the too-often disparate disciplines of applied ethics and health economics, namely with a view to benefitting each.

Careful attention to the work of some health economists on how to conceptualise and operationalise ‘need’ as an allocative principle can serve to advance current debates over allocative principles. Those debates span health economics, ethics, policy, and health technology assessment (the field of inquiry in which the costs, effects and broader implications of health interventions are systematically

evaluated). The debates often centre on how an orthodoxy of contemporary health economics and policy ought to be revised, or at least tempered or supplemented. That orthodoxy comes to a single allocative principle, namely ‘allocate resources according to expected benefits’ or, much more precisely, ‘allocate resources so as to maximise aggregate health’. Revision of this orthodoxy is often urged in the name of need, something to which the orthodoxy is deemed to inadequately answer, morally speaking. A recent review conducted by a team of health economists and ethicists for the World Health Organisation (Norheim et al. 2014) identified numerous factors that orthodox economic analysis fails to attend to but which nonetheless can be relevant in assessing the value of a health intervention. Many of these factors can be conceptualised in terms of need. Toward the end of this article, we examine two government-backed efforts to temper the orthodoxy of maximising aggregate health by operationalising some notion of ‘need’ as an allocative principle.

Cookson and Dolan (2000) have expounded a number of competing allocative principles and used a hypothetical scenario to illustrate the often divergent allocations of health care resources that these commend. In the first half of this article, we do likewise. However, our hypothetical scenario is left quite abstract, so as to keep the reader’s focus on the general issues. Furthermore, our comparison of allocative principles is sharpened by two analytical tools. First, using quantification helps to clarify divergent consequences by bringing a degree of precision (Williams 1997, 120). Second, we employ a diagrammatic approach developed by Culyer (2006) and Wagstaff (Culyer and Wagstaff 1993), because it offers a visual and conceptual aid.

Our comparison draws out how, of the eight principles we illustrate, only two principles treat as relevant both a person’s initial health state and a person’s capacity to benefit per resource unit expended. Those two allocative principles are: (1) allocate

resources so as to most closely equalise final health states, and (2) allocate resources so as to equally restore health states to population norms. These allocative principles ought to be preferred to alternatives if one takes seriously the relevance of both a person's initial health state and a person's capacity to benefit per resource unit expended. The contemporary orthodoxy of 'allocate resources so as to maximise aggregate health' is limited in that it treats as relevant solely a person's capacity to benefit per resource unit expended, as we go on to demonstrate.

Our comparison of allocative principles is also intended to provide insights and a potential framework for use in any deliberative process that might be undertaken to establish how health care resources ought to be allocated. Daniels (1993, 1994; 2009, 38) continues to write of "unsolved rationing problems." These are problems that theoretical debate has failed to solve over the course of decades, problems over which reasonable people disagree and which seem most easily (perhaps even solely) solved through a fair procedure taking the form of a deliberative process, like a Citizens' Jury (Smith and Wales 2000).<sup>1</sup> "Establishing a fair process for priority setting is easier than agreeing on principles", Daniels (2000) writes. While this may be true, principles are nonetheless "important inputs into a fair, deliberative process" (Daniels 2001, 10). Meaningful deliberation can never be all process. In this vein, deliberations about health care resource allocation can greatly benefit from the kind of comparison of allocative principles that we undertake in this article.

Early on, we confine our analysis to four assumptions that are contestable but nonetheless common in the practice of health economics and policy. (1) Resources are scarce relative to the totality of worthy possible uses. (2) As such, resources should be allocated justly. (3) Justice is to be judged purely by reference to specified consequences for health or, in other words, intervention effects. (4) Every uniform

increment or “quantum” of intervention effect is assumed to be of equal value (Williams 1997, 120). (In other words, each constitutes an equal benefit.) Later in the article, we examine departures from the fourth assumption, especially.<sup>2</sup>

Following our illustration, we examine the possibility of balancing multiple allocative principles. When seeking to balance health maximising and health equalising principles, the question of their relative importance can be seen as a question concerning the rate at which one is willing to trade aggregate health for greater health equality. We then analyse proposals concerning the importance of securing for all a ‘fair innings’ (Williams 1997). These proposals encompass a lifetime view of health, such that, for the allocator of resources, health is to be viewed less as a state than as a kind of stock that accumulates over time. We then demonstrate how uniform quanta of intervention effect can be differentially valued, variously in the name of securing for all a fair innings, in the name of prioritising the worst off, and in the name of need. (These can all come to one and the same.) By doing this, we examine departures from the fourth assumption above. Finally, we examine government-backed efforts in the Netherlands and Britain to differentially value quanta of effect as a means of incorporating some notion of ‘need’ into the (routine, systematic, quantitative) valuation of health interventions. Both efforts represent one type of Prioritarianism (a prioritising of the worst off), and we demonstrate precisely how this is so, for the first time differentiating between different types of Prioritarianism. We thus draw attention to burgeoning efforts within health economics and policy circles to operationalise insights made by philosophers, among others, concerning the importance of more than simply maximising aggregate health. We do this while simultaneously inviting philosophers to benefit from the conceptual resources developed by health economists and earlier demonstrated in our illustration.

## II. ALLOCATIVE PRINCIPLES

A principle can simply be a guide to action or rule of thumb (Oxford English Dictionary 2007). Aristotle's formal principle of justice states that equals should be treated equally and, by corollary, unequals should be treated unequally (Beauchamp and Childress 2009, 242). That principle is merely formal insofar as the *substance* of what makes for equals is not yet specified. "Substantive" or "material" principles of justice specify what equals share (Beauchamp and Childress 2009, 242). They specify *relevant respects* in which people or their claims are equivalent. Equals are equal in their *X*. *X* is precisely a substantive principle of justice. While Aristotle's formal principle does not identify equals, it does specify that, once identified, equals should be *treated* equally. Its accent falls on treatment, not on outcome. Equal treatment, whatever it is, may or may not include ensuring that equals enjoy equal outcomes.

The first substantive principles of justice that we differentiate can be formulated as follows. Scarce health care resources ought to be allocated in direct proportion to each person's:

1. incremental capacity to benefit
2. total capacity to benefit
3. need, considered as inversely related to their initial health state
4. need, considered as the resources required to exhaust all capacity to benefit

(Culyer 2001, 2007)

A person's 'incremental capacity to benefit' refers to a person's capacity to benefit per resource unit expended. (This can change, usually diminishing, as more and more resources are expended.) In this article, we often speak of persons, but substantive principles of justice can also be applied in relation to relevantly homogenous groups. In this way, the word 'parties' can be used to encompass both persons and groups.

We then examine the following principles of justice. They represent *goals* for health care resource allocation, and can be formulated as follows. Scarce health care resources ought to be allocated so as to:

5. maximise aggregate health
6. equally exhaust each person's capacity to benefit
7. most closely equalise final health states
8. equally restore health states to population norms

Aggregate health is maximised by allocating resources in direct proportion to each person's incremental capacity to benefit – we demonstrate this below. Principles 5 and 1 are therefore equivalent; each represents merely a reformulation of the other.

Further allocative principles are conceivable. For example, scarce health care resources could be allocated in direct proportion to each person's *willingness to pay*. However, we confine our analysis to Principles 2-8 as formulated above, partly because each is illuminated by comparison with the others using the analytical tools that we adopt, namely the quantification of health states and the diagrammatic approach developed by Culyer and Wagstaff.<sup>3</sup>

### III. ILLUSTRATION



The intervention will form our basic resource or cost unit. Imagine a standardised treatment, like a course of medication or therapy, whose financial and other costs remain constant but whose therapeutic benefit tends to diminish as a person receives successive treatments. The intervention can be equated to an allotment of public funding; in this way, a fraction of an intervention can be intuited as a fraction of the funding needed to cover the cost of the intervention. The rest of the intervention could be paid for privately, for instance.<sup>4</sup>

Our measure of benefit will be health units (h), an abstract, hypothetical unit able to track changes in the health of a person. Culyer (2006, 1156) puts it thus: “Suppose that ... health, is measured on some linear scale, which is agreed to have construct validity, which allows the measurement of more or less of it.” Imagine a unit like metres that a person can see, degrees that a person can move their arm, or something more abstract like the quality of life routinely quantified using an instrument such as the EQ-5D (Kind, Brooks, and Rabin 2005).<sup>5</sup>

Given the above, let us suppose the following.

- Six interventions are to be immediately allocated between two persons.
- Both persons share the same initial health state: without an intervention, their health state will remain at zero. Hence, in Figure 1, this starting point, or “endowment point”, coincides with the origin (Wagstaff 1991, 24).<sup>6</sup>
- Conversely, each person’s health state will tend to increase with every additional intervention, but differently for each person and in non-linear fashion: successive interventions will tend to increase the health state less and less, that is, in diminishing increments. Figure 1 plots the different health states achievable by different allocations of all available resources.

At the point marked (0,6), A is allocated zero interventions while B is allocated all six. As a consequence, A's health remains unchanged at zero while B's health increases from zero to 23h. Likewise, at the point marked (1,5), A is allocated one intervention while B is allocated the remaining five. As a consequence, A's health increases from zero to 33h while B's health increases from zero to 22h. And so on.

How ought the six interventions to be allocated? The answer depends on which allocative principle is favoured.

### Plotting the health frontier

The kind of curve plotted in Figure 1 has been called "the health frontier" (Wagstaff 1991, 24). It represents the range of desirable, attainable states for any two parties under a particular resource constraint (the availability of six interventions, in our illustration). In this way, allocating constrained resources between two parties is equivalent to selecting a point along the health frontier.

The frontier's shape is defined by the constraint on resources together with the parties' respective capacities to benefit from them (in other words, the respective rates at which parties can convert resources into benefits). Increasing these rates, or lifting the resource constraint, will push the frontier further out from the endowment point. In practical terms, this could take the form of improving intervention effectiveness, or increasing the number of available resources.

By definition, every point along the health frontier is Pareto-efficient. This means that no further 'free' benefits are available. All resource waste has been eliminated, for instance. *Benefits accruing to one party can now only be achieved with*

*the loss of benefits to another.* We must trade off one party's benefits for another's. The rate of trade is indicated by the slope of the curve at any one point. When the curve section is perfectly horizontal or vertical, the cost to one party of benefitting the other is zero; that is, a 'free' benefit is available. As such, perfectly horizontal and vertical sections are not properly a part of the health frontier. For instance, in Figure 1, the curve section from (5,1) to (6,0) is perfectly vertical. This signifies that A's health is not improved by a sixth intervention. That intervention can therefore be allocated to B with no loss to A in terms of benefits denied. This curve section is thus excluded from the health frontier proper. Indeed, Culyer (1992; 2006, 1155) argues that it would be unethical to entertain allocations that forego 'free' benefits. Such a concern for Pareto-efficiency can be considered a stand-alone ethical principle. It is maximising (of 'free' benefits) and underpins any use of the health frontier as a conceptual aid to allocative thinking.

The next sections examine the application of Principles 2-8 in the context of our illustration.

Principle 5 (= Principle 1). Allocate resources so as to maximise aggregate health

While every point along the health frontier is Pareto-efficient, not every point maximises aggregate health. Aggregate health is maximised only by that point intersecting with the -1 slope. On either side of this point, a resource unit attains benefits for one party smaller than those that it could otherwise attain for the other party, thus failing to maximise aggregate health. At the point along the health frontier that intersects with the -1 slope, benefits accrue to one party with equivalent loss to the other in terms of benefits denied (that is, at a 1:1 ratio). Put another way, parties now

share an equivalent incremental capacity to benefit (Principle 1). This is why Principle 5 equates to Principle 1. In Figure 1, we can see that the curve section from (3, 3) to (4, 2) will contain a point that intersects with the -1 slope. Making assumptions about the curve's shape, namely its mathematical function, enables one to identify that point exactly.

Principle 5 is commonly criticised for virtually ignoring initial health states. Initial health states are treated as irrelevant beyond being needed to calculate the magnitude of the first increments of benefit. On Principle 5, solely a person's incremental capacity to benefit is relevant; if this is already known, then one need not inquire further into the person's initial health state. As a result, resources can be allocated to healthier people so long as those people exhibit greater incremental capacities to benefit. This offends against moral intuitions that oppose benefitting the better off at some cost to the worse off. Put differently, it offends against moral intuitions that commend prioritising those most in need.

Principle 2. Allocate resources in direct proportion to each person's total capacity to benefit

A person's capacity to benefit can be considered not only in per-resource-unit increments but also *in toto*. This furnishes a very different substantive principle of justice. A person's total capacity to benefit is a kind of theoretical maximum defined by the effectiveness of current interventions and considered independently of resource constraints. Figure 1 indicates that A's total capacity to benefit corresponds to an increase of 65h. This is the greatest benefit that A can hope to attain given the effectiveness of current interventions. As explained above, the perfect verticality of

the health frontier from (5, 1) to (6, 0) indicates that A does not benefit from a sixth intervention and, by extrapolation, from any subsequent intervention. By contrast, the slope of the health frontier does not reach horizontal. As such, B still stands to benefit (presumably in diminishing increments) from interventions beyond the sixth, that is, from resources beyond those currently available. Making some assumptions about the shape of the curve, we can estimate that B's total capacity to benefit equates to an increase of 24h.

Principle 2 can be criticised on at least two scores. It treats as irrelevant incremental capacities to benefit. As a result, it risks foregoing larger benefits that are otherwise attainable. Furthermore, like Principle 5, Principle 2 treats as irrelevant initial health states, beyond needing them to calculate capacities to benefit. As such, like Principle 5, it can favour those already better off.

Principle 3. Allocate resources in direct proportion to each person's need, considered as inversely related to their initial health state

A person's need is sometimes conceptualised as his or her capacity to benefit, whether incremental or total (Hasman, Hope, and Østerdal 2006; Hope, Osterdal, and Hasman 2010). In this case, Principles 2 and 5 remain unchanged, but appear under a different name: 'need' simply replaces 'capacity to benefit.'

A person's need can be conceptualised otherwise, namely as inversely related to his or her initial health state (Hasman, Hope, and Østerdal 2006; Hope, Osterdal, and Hasman 2010). The less the initial health state, the greater the need.<sup>7</sup> On this conception, people who share the same initial health state share the same degree of need. In our illustration, both persons share the same initial health state, namely zero

h. Principle 3 therefore commends an equal allocation of resources for each person: three interventions each. It does this irrespective of outcome.

Principle 3 can be criticised for treating as irrelevant capacities to benefit, both incremental and total. As a result, it is blind to what benefits might be foregone and, in this respect, is potentially wasteful of resources.

Principle 4. Allocate resources in direct proportion to each person's need, considered as the resources required to exhaust their total capacity to benefit

Culyer (2001, 2007) advocates a different conception of need, though he does not favour it as an allocative principle. For Culyer, a person's need – for a specific form of health care – *derives* from their total capacity to benefit from that health care. It does not *equate* to that capacity, however: “need is defined as the resources, valued in expenditure terms, required to exhaust capacity to benefit” (Culyer 2001, 279).

On this conception, A needs five interventions, since A's total capacity to benefit is exhausted after five interventions. B's total capacity to benefit is not exhausted after five interventions, so B needs more interventions than A, on Culyer's conception of need. Principle 4 then commends proportional allocation of the available resources. For instance, if B happens to need 12 interventions, while A needs 5, then of the available six interventions, B will be allocated 4.2 and A will be allocated 1.8.<sup>8</sup>

Principle 4 is blind to diminishing increments of benefit. In this vein, Culyer (2007, 237) observes: “it is not immediately apparent that this is an attractive [allocative] criterion. It will favour those whose marginal capacity to benefit falls to zero only after very large expenditures have been incurred.”

Principle 6. Allocate resources so as to equally exhaust each person's total capacity to benefit

As per Principle 4, applying Principle 6 requires estimation. This time, we must estimate B's total capacity to benefit rather than the number of interventions required to exhaust it (which we imagined to be 12, for argument's sake). Looking at the health frontier and making particular assumptions about its shape, we can estimate that B's total capacity to benefit is exhausted after accruing 24h. As noted above, A's total capacity to benefit is exhausted after accruing 65h. In Figure 1, the point 'Total CTB' marks each person's total capacity to benefit (according to our estimation). We can draw a line from this point to the endowment point. The point at which this line intersects with the health frontier indicates the allocation required to equally exhaust each person's total capacity to benefit, preserving the relative difference between them. That relative difference is indicated by the slope of the 'Total CTB' line ( $24h/65h = 0.37$ ). In Figure 1, the commended allocation (corresponding to the point of intersection) is approximately (2.5, 3.5). This allocation results in equally exhausting each person's total capacity to benefit, namely by approximately 83%.<sup>9</sup>

Principle 6 reflects one view expressed by the public, namely that “‘realizing one's potential for health improvement,’ whether large or small, [is] important in social decision-making” (Beauchamp and Childress 2009, 572). Nonetheless, Principle 6 can be criticised on precisely the same scores as Principle 2. In treating as irrelevant incremental capacities to benefit, Principle 6 risks foregoing larger benefits that are otherwise attainable. Moreover, in treating as irrelevant initial health states,

beyond needing them to calculate capacities to benefit, Principle 6 can favour those who are already better off.

Principle 7. Allocate resources so as to most closely equalise final health states

Culyer (2001, 2007) argues that health care allocation should aim, not to maximise aggregate health, but to *minimise inequalities in health*. Allocation should aim to equalise people's health as far as this is allowed by the health frontier, that is, by available resources, levels of intervention effectiveness, and a concern for Pareto-efficiency.

In Figure 1, the 45 degree line (marked 'slope = 1') identifies all possible points at which the health states of A and B would be equal. Only the point on the health frontier that intersects with this line identifies the allocation that will equalise final health states. In Figure 1, this is when A and B both have 22h. Solely the allocation that approximates to 0.7 interventions for A and 5.3 interventions for B will equalise final health states.<sup>10</sup>

In Figure 1, the health frontier would not have intersected with the 45 degree line if one person's initial health state had already exceeded the best possible health state attainable by the other person (given intervention effectiveness and available resources). In this case, it would not have been possible to equalise health states, without reducing one person's health. Culyer (2001, 281) argues that allocation should aim to equalise health "probably" without reducing any identifiable individual's health. Principle 7 then commends the allocation corresponding to the point on the health frontier that is nearest to the 45 degree line.



Principle 7 treats as relevant both initial health states and incremental capacities to benefits. However, aggregate benefits will often be foregone in the interests of most closely equalising final health states.

Principle 8. Allocate resources so as to equally restore health states to population norms

Let us articulate Culyer's view more precisely. Culyer does *not* argue that allocation should aim to equalise the health of *all* persons. More precisely, Culyer (2006, 1156–1157) argues that allocation should aim to equalise the health of all persons who are *relevantly similar* or, in his words, “ethically homogenous.” The familiar question then arises as to precisely which persons are relevantly similar: who are equals?

Principle 7 implies that *all* persons are relevantly similar (equals) insofar as allocation should aim to equalise their health. Equals can be conceptualised differently, however. For instance, we could regard as equals those whose initial health states compare equally *relative to the normal* health state of the relevant cohort or population.<sup>11</sup> My health might be much better than yours, but we might both be only half as healthy as our respective peers. In this way, we can be considered equals. Culyer (2001, 281) raises this possibility, identifying relevant populations (peers) by age or, more accurately, “stage of the life-cycle”:

An equitable health care policy should seek to reduce the inequality in health (life expectation, self-reported morbidity, quality of life in terms of personal and social functioning) at every stage of the life-cycle. Such a

policy must meet needs, but in proportion to the ‘distance’ each individual is from the population average.

In similar fashion, Daniels conceptualises health in terms of normal or, more precisely, species-typical functioning, and insofar as this declines with age, health care policy ought to protect or restore each person’s “age-relative opportunity range” (2008, 478, 481).

Principles 7 and 8 differ in intended outcomes. While Principle 7 reflects an intention to improve health, Principle 8 reflects an intention to improve health only in line with what is normal for a population. For instance, female fertility normally declines with age. Returning to Figure 1, imagine that A is younger than B and that, as such, the fertility levels normal for their ages differ. Imagine that these levels correspond to 66h and 44h for A and B, respectively.<sup>12</sup> Point ‘Population average’ marks this as the intended outcome. As per Principle 6, we can draw a line from this point to the endowment point. The point at which this line intersects with the health frontier – namely (1, 5) – indicates the allocation required to equally restore health states to population norms. The allocation of (1, 5) happens to restore the health states of both persons to exactly half the respective population averages. It preserves the relative difference between those averages, indicated by the slope of the ‘Population average’ line ( $44h/66h = 0.67$ ). Principle 8 implies that, with intervention, B’s fertility *ought to be* only 0.67 times that of A’s fertility, in line with population averages.

Points ‘Total CTB’ and ‘Population average’ both represent intended outcomes. The first point is located by the (maximum, cumulative) effectiveness of available interventions. It lies beyond the health frontier because the intended outcome that it marks outstrips available resources. By contrast, point ‘Population

average' lies beyond the health frontier because the intended outcome it marks outstrips, not available resources, but the *effectiveness* of available interventions. Availing more resources will push the health frontier out from the origin but only ever to the point 'Total CTB'. In the case of Figure 1, the health frontier will never reach the 'Population average' point. This is due to current intervention effectiveness. For A, the vertical curve section between (5, 1) and (6, 0) represents a limit that additional resources cannot overcome. Likewise, for B, the curve will asymptote horizontally well before being pushed up to 44h.

Like Principle 7, Principle 8 commends an allocation heavily in favour of B. The allocation of (1, 5) opposes a double whammy. Williams (1997, 118) and others (Savulescu 1998, 215) write of a "double jeopardy" or "double injustice" when, to maximise aggregate health, health care is diverted from people with low health states by virtue of their also having small capacities to benefit from current interventions. The allocation of (1, 5) compensates for the diminished *effectiveness* of interventions for B. But it does not, *and in principle does not intend to*, compensate for diminished fertility insofar as this normally declines with age. In this way, Principle 8 contrasts with Principle 7. Principle 7 *does* intend to compensate for (to overcome) declines *both* in intervention effectiveness *and* in normal fertility. Meanwhile, Principle 5 reallocates resources away from B in view of the first decline, redoubling what B suffers in view of the second decline. Principle 5 is concerned with B's initial health state only insofar as that state bears influence on the magnitude of B's first increment of benefit. In stark contrast to Principle 5, Principles 7 and 8 allocate interventions to B rather than A precisely because the interventions are *less effective* at increasing health (in absolute terms and relative to population averages, respectively). In a word, B *needs* more interventions, because interventions for B are less effective.

Principle 8 is subject to much the same *prima facie* assessment as Principle 7. It treats as relevant both initial health states and incremental capacities to benefits, but also foregoes some aggregate benefit.

Principle 8 can be understood as reflecting an appeal to nature, if population norms are identified with nature. Principle 8 is potentially consistent with a conception of medicine on which the proper role of medicine is to restore nature but not to improve on it (Carter and Braunack-Mayer 2011). By contrast, Principle 7 is compatible with a conception of medicine on which medicine is not morally constrained by whatever is normal or natural. This can be seen as an advantage or as a limitation, depending on one's moral's view. One needs to confront fundamental questions concerning the proper role of medicine if one is to adjudicate between Principles 7 and 8.<sup>13</sup> Put differently, one needs to form a view within the bioethics 'enhancement' debate.

Our illustration demonstrates that competing allocative principles can commend divergent allocations, with similarly divergent consequences for health. As fewer interventions are allocated to A and more to B, the difference between their final health states tends to diminish but so too does the sum of those states (aggregate health). Under Principle 8, some difference in the final health states of A and B is, far from being unwelcome, precisely the outcome that is intended: the difference should persist in line with population averages.

For the rest of this article, we move beyond our illustration to explore possibilities for balancing multiple allocative principles and for conceptualising benefits as relative to one's initial health.

#### IV. THE RATE AT WHICH ONE MAY BE WILLING TO TRADE AGGREGATE HEALTH FOR HEALTH EQUALITY

Principle 5 can tend to dominate the contemporary practice of health economics and policy. However, Principle 5 is merely one of a number of competing allocative principles, some of which are better poised to reduce rather than to exacerbate differences in final health states. There is increasing empirical evidence that the public does not favour the exclusive use of Principle 5, preferring to sacrifice aggregate health gain for greater health equality (Richardson et al. 2012). A review of 64 empirical studies found that the public tends to value more highly those health gains that help to reduce health inequalities (Dolan et al. 2005). When given the choice between two identical health gains, the public strongly prefers the gain going to the person with the poorest health. This preference has been independently observed in nine studies conducted across Australia, Canada, Norway, Spain and the USA (McKie and Richardson 2005).

A concern for justice can encompass concerns both to maximise aggregate health and to minimise health inequality. However, these concerns will often compete. Therefore, at what rate, if any, are we willing to trade aggregate health for greater health equality?

This trade-off is often described as being between ‘efficiency’ and ‘equity.’ However, this description is potentially misleading, for both of these terms can be given very different meanings. Efficiency can be identified with Principle 5 or, more generally, with Pareto-efficiency and, in this respect, with plotting the health frontier. Likewise, equity can be identified with, for one, Principle 7 or, more generally, with justice in resource allocation (and thus with choosing a point along the health frontier).

Culyer (2006) argues that ‘efficiency’ and ‘equity’ ought to be given the latter meanings and that, as such, the conflict between them is “bogus” – there is none. We plot the health frontier then choose a point along it. There remains, however, a genuine conflict between Principle 5 and, for one, Principle 7, as reflected in the distance between the points that they respectively lead us to choose along the health frontier in Figure 1.

Someone who balances competing commitments to Principles 5 and 7 is, in Parfit’s terms, a *pluralist, moderate Telic egalitarian*. By ‘Telic,’ Parfit means consequentialist. A Telic egalitarian values equality with respect to consequences, e.g. final health states. A *moderate* Telic egalitarian is prepared to trade some equality for utility (e.g. aggregate health) provided that no one is made worse off. For instance, improving one person’s health but not another’s strictly reduces equality, but no one is made worse off with respect to health states. This simply equates to a concern for Pareto-efficiency. A *pluralist* Telic egalitarian is prepared to trade some equality for utility (e.g. aggregate health) *even when someone is made worse off*, provided that the trade is worth it.

If we can quantify the rate at which we are willing to trade aggregate health for greater health equality, then we can use this number to define the shape of a “social welfare function” (Wagstaff 1991, 123–125; Williams 1997). In Figure 1, this would be a curve that reflects the rate at which we are willing to trade. It would reflect the relative strength of our commitments to Principles 5 and 7. The curve would intersect with the health frontier at some point between the points where the -1 slope (Principle 5) and the 1 slope (Principle 7) intersect with the health frontier.

A social welfare function could instead be used to reflect the relative strength of commitments to Principles 5 and 8. In this case, maximising aggregate health is

balanced against equalising the health, not of all persons, but of those relevantly similar, now identified with reference to population norms. Aggregate health is traded with the health equality of peers so identified.

## V. A FAIR INNINGS

Principle 8 can instead be understood as reflecting an appeal, not to nature, but to a fair innings. These appeals are different and distinct, though they can be run together, e.g. we can appeal to nature as establishing whatever constitutes a fair innings.

Williams (1997) appeals to a fair innings as the outcome that allocation should intend for each person, considering health as a kind of stock that accumulates over time in increments, with the magnitude of those increments depending on one's health-related quality of life. Williams (1997, 119) considers the appeal to a fair innings to be (1) consequentialist, (2) "about a person's *whole life-time experience*, not about their state at any particular point in time",<sup>14</sup> and (3) reflective of an "aversion to inequality" (that is, broadly egalitarian). Moreover, the fair innings itself is (4) quantifiable. We should all enjoy a similar, if not equal, quantity of health over a lifetime. (In this light, a question arises concerning our illustration: what additional health stock has B already accumulated, if B is older?)

There is evidence of public support for a view like Williams's. McKie and Richardson (2005, 3) flag a widely shared "preference for younger patients because of an aversion to inequality in age of death", or in final accumulated health stock, to use Williams's terms. Surveys of the Australian public have found a predominant concern to equalise individuals' lifetime health prospects (Richardson et al. 2012). This

preference has been called “egalitarian ageism”, since it will often favour the young (McKie and Richardson 2005, 3).<sup>15</sup>

An advocate of a fair innings for all is only interested in age insofar as age proxies for an accumulated stock of health. Strictly speaking, age is morally irrelevant for the fair-innings advocate; it is the accumulated stock of health that is relevant: “the justification is not age itself, but who the worst off are in terms of lifetime health” (Norheim et al. 2014, 6). For instance, if one person is old but has suffered an incredibly poor quality of life all along, then their claim on health resources can still be strong, indeed stronger than that of a much younger person who has enjoyed perfect health, for the younger person has had, in a word, a better innings (having accumulated a greater stock of health-adjusted life years, for instance).

A fair innings can be taken to mean, not an equal stock, nor even a comparable stock, but simply a reasonable stock. Frankfurt (1987, 21) writes of, not health, but wealth: “If everyone had enough, it would be of no moral consequence whether some had more than others.” Perhaps allocators are particularly willing to trade aggregate health for greater health equality when one party does not have ‘enough’. If so, then the rate at which we are willing to trade is not constant: it changes depending on how poor is someone’s health, or how low is their stock. Put differently, the relevance of a person’s capacity to benefit changes depending on how poor is their initial health, or how low is their stock. Here, equal increments or quanta of intervention effect are no longer *valued* equally: they no longer constitute the same benefit. We return to this issue below, particularly in discussing an allocative principle proposed by Bognar (2015).

Norheim et al. (2014) repeatedly assert that all persons should have “a fair chance to live a full and healthy life.” Williams could well agree with this, but in his



particular advocacy of a fair innings for all, he would sharpen the assertion as follows. All persons should have an *equal probability* to enjoy an *equal stock* of health-adjusted life years. This assertion contrasts quite radically with the following. All persons should have an *equal opportunity* to live *the whole of their natural lives*. This may be closer to what Norheim et al. mean. A concern to secure for each person a fair innings is precisely a concern to secure for each person some threshold or equivalent stock of health, not to secure for each person the whole of their natural life.<sup>16</sup> It is for this reason that, on its own, the concern to secure for each person a fair innings is inadequate to protect older people, insofar as those deemed to have already had a fair innings (the requisite stock) are deemed to have no justice-based claim on health resources.

For all the conceptual resources that health economists can provide ethicists – and we hope to have demonstrated the power of some of them through our illustration – health economists may, for their part, nonetheless need to wrestle with the moral importance, not merely of probabilities but of opportunities, and not merely of added quanta (stocks) of health, but of people’s whole natural lives.

## VI. PRIORITARIANS

Parfit (1997) coined the term ‘Prioritarian’. A Prioritarian *prioritises* the worst off, independent of concerns for equality or aggregate utility.<sup>17</sup> Prioritising the worst off out of duty or out of a concern for consequences makes you either a Deontic or Telic Prioritarian. Prioritarianism is consonant (though not equivalent) with wanting parties to have “enough”, that is, with valuing “sufficiency” above equality (Frankfurt 1987,

22). Parfit (1997, 212) expresses the Prioritarian view in a number of ways, most simply thus: “we might be especially concerned about those people who are worse off.” (St Paul was asked to “remember the poor” (Galatians 2:10).) Parfit (1997, 221) also expresses the Prioritarian view thus: “gains to the worse off count for more.” Parfit (1997, 213) formally defines the Prioritarian view thus: “Benefiting people matters more the worse off these people are.”

On close examination, these different expressions contain at least two distinct views, which Parfit failed to differentiate. One type of Prioritarian simply prioritises the worst off, while a second type is concerned with benefits considered relative to how badly off people are to begin with. For instance, suppose that P has 9 while Q has 10, and we can give 1 to P or 10 to Q. The first type of Prioritarian will give 1 to P, since P has the least: they simply prioritise the worst off. The second type of Prioritarian *might* give 10 to Q because  $10/10 \gg 1/9$ ; that is, the second type of Prioritarian seeks to maximise benefits considered relative to how badly off people are to begin with. We say ‘might’ because, for the second type of Prioritarian, a question remains as to *how much* “gains to the worse off count for more” (Parfit 1997, 221). In other words, what is *the rate* at which we are willing to trade benefits going to the better off for benefits going to the worse off? The first type of Prioritarian does not ask at what rate gains accruing to one party should be traded for gains accruing to another. Instead, she simply allocates resources to the worst off. For second-type Prioritarians, uniform increments or quanta of intervention effect are no longer valued uniformly. Some are valued more than others; some are taken to constitute greater benefits than others. Benefits are redefined as *relative* to how badly off someone is, e.g. in terms of pre-intervention health. Second-type Prioritarians can seek to

maximise such benefits in the aggregate, or they can seek to equalise such benefits among parties. We now go on to demonstrate this.

Bognar (2015) endeavours to discern an allocative principle that conforms to, and might therein be seen to underpin, a series of moral intuitions regarding which of two persons ought to be allocated a resource unit, when those persons differ only in age and/or the capacity to benefit. Bognar (2015) intuites the following.

- A “20-year old patient who will live for many years if she gets the drug” ought to get the drug above “a 70-year old patient who will live for only a few more years if she gets the drug.”
- A “20-year old patient who will live for 10 more years if she gets the drug” ought to get the drug above “a 70-year old patient who will live for 10 more years if she gets the drug.”
- A “30-year old patient who will live for 10 more years if she gets the drug” ought to get the drug above “a 40-year old patient who will live for 10 more years if she gets the drug”.<sup>18</sup>
- A “30-year old patient who will live for 10 more years if she gets the drug” ought to get the drug above “a 60-year old patient who will live for 10 more years if she gets the drug”.<sup>19</sup>

Bognar proposes that the best allocative principle conforming to these intuitions is as follows: allocate resources so as to maximise, in the aggregate, life years whose assigned value diminishes with age. In other words, allocate resources so as to maximise aggregated benefits, where quanta of effect (i.e. life years) constitute ever smaller benefits as they accrue to an individual. Bognar (2015, 10) envisages a case of diminishing marginal returns: “we should assign different weights to additional life

years, but their value should be diminishing as they are added to the lives of older persons.”

Bognar characterises his view as Prioritarian. More precisely, his view accords with what we have called *second-type Prioritarianism*, for on Bognar’s view, it is conceptually possible to specify a rate of trade that identifies precisely *how much* more valuable are quanta of effect (i.e. life years) accruing to the people worst off (i.e. the youngest or, more precisely, those having so far accumulated the least health stock). Moreover, Bognar is a second-type Prioritarian *of the maximising kind*, since he seeks to maximise aggregated benefits, where benefits comprise quanta of effect assigned a value relative to how badly off someone is. But there is an alternative allocative principle that Bognar does not explicitly discuss but which equally commends allocations conforming to his intuitions, namely *first-type Prioritarianism*. To reiterate, the first-type Prioritarian does not ask at what rate gains accruing to the better off should be traded for gains accruing to the worse off. Instead, she simply allocates resources to the worst off; other considerations are trumped. For example, the first-type Prioritarian *always* secures an extra year of life for a 20-year old, no matter how many life years she could otherwise secure for a 70-year old, and no matter how many 70-year olds she could otherwise secure an extra life year for. Bognar suggests that this is precisely why his allocative principle is superior: “it allows trade-offs to be made between benefits to different people.” However, Bognar is left with “the aggregation problem”, namely the problem of how much we ought to prioritise large gains for the few over modest gains for the many (Daniels 1993). Bognar (2015, 10) simply requires that lots and lots of minor health benefits (“*many more*”) be traded for somebody’s life. By contrast, the first-type Prioritarian is not left with the aggregation problem. Nor are Williams and Norheim et al.: they treat people,

not quanta of health, as moral equals, insofar as they assert the following. All persons should have an equal probability to enjoy an equal stock of health-adjusted life years (Williams). All persons should have an equal opportunity to live the whole of their natural lives (Norheim et al.).

Differentially valuing quanta of effect in favour of the worst off: recent Dutch and British efforts at operationalising need as an allocative principle

In effect, Bognar proposed one way in which we might choose to differentially value uniform quanta of effect (i.e. life years) in the manner of a second-type Prioritarian, namely prioritising the worst off at a particular rate of trade (put differently, according to a particular mathematical function). Differentially valuing quanta of effect in such a way can be undertaken in the name of need. That is, it can constitute an attempt at operationalising need as an allocative principle. We now demonstrate this by explicating, by the light of some of our preceding analysis, two efforts made by Dutch and British health economists to operationalise need as an allocative principle.

Since at least 1991, the Dutch have formally acknowledged the importance of a health intervention's "effectiveness", "efficiency" and "necessity" (van de Wetering et al. 2013, 108). In the Netherlands, as in comparable countries, efficiency has increasingly been operationalised as an allocative principle by the mandating of formal economic evaluations, which typically compare an intervention's costs and effects to those of alternative interventions with a view to maximising aggregate health (Principle 5). Meanwhile, necessity

while intuitively important, eluded definition and measurement and thus remained barely applied as a (systematic) selection criterion ...

Proportional shortfall was introduced to provide a more systematic and quantitative definition of necessity (van de Wetering et al. 2013, 111).

Proportional shortfall refers to the proportion of remaining lifetime health that one stands to lose by virtue of an illness, assuming that the illness goes untreated. It refers to how far someone stands to *fall short* of a full and healthy life in the absence of treatment, expressed as a *proportion* of the remaining health the person could otherwise expect in the absence of illness. As an allocative principle, proportional shortfall embodies the morally normative view that “priority should be given to those who lose the greatest *proportion* of their remaining health expectancy due to some illness if the illness remains untreated” (van de Wetering et al. 2013, 110).

Van de Wetering et al. quantify proportional shortfall with the following equation, which features the orthodox unit of measure for a health gain or loss, the quality-adjusted life year (QALY).<sup>20</sup>

Disease-related QALY loss

---

Proportional shortfall =  $\frac{\text{Disease-related QALY loss}}{\text{Remaining QALY expectation in absence of the disease}}$

This equation can be unpacked (see Figure 2). Van de Wetering et al. (2013, 110) note that “the denominator reflects the remaining QALY expectation in normal health, which could, for example, be determined on the basis of the age and gender.” Moral decisions are unavoidably involved “in defining the relevant group”, namely in determining who ought to be considered a person’s peers and therefore what normal

health comprises (precisely as per Principle 8) (van de Wetering et al. 2013, 112). Van de Wetering et al. (2013, 110) explain that the “proportional shortfall is 1 for all patients who face a threat of immediate death, irrespective of their age”, since “they will lose 100% of their remaining life.” Likewise, the proportional shortfall is 0.5 for all persons who stand to lose half of the QALYs they could otherwise expect to accrue, irrespective of how many QALYs this is. In this way, the younger person who stands to lose 20 of 40 expected QALYs, and the older person who stands to lose 1 of 2 expected QALYs, are attributed the same necessity, or degree of need.

Van de Wetering et al. (2013, 111) remark that “proportional shortfall proposes to equalize relative attainments.” This suggests the following allocative principle: allocate resources so as to recover for each person the same proportion of health (e.g. half the QALYs) that would otherwise be lost. This principle conforms to second-type Prioritarianism *of the equalising kind*. The intended outcome is to equalise benefits among parties, where benefits comprise quanta of effect (e.g. QALYs) differentially valued according to how badly off someone is. In the case of proportional shortfall, the greater the proportion of one’s remaining life one stands to lose, the worse off one is. (More precisely, if we return to Figure 2, the bigger A is relative to B, the worse off someone is.) However, van de Wetering et al.’s (2013, 111) observation that proportional shortfall proposes to equalise relative attainments seems mistaken, for on its own, proportional shortfall has nothing to do with *attainments*, namely quanta of treatment effect. Proportional shortfall has only to do with expected *losses*, namely by virtue of illness. What can be done about the illness has yet to enter the picture.

Van de Wetering et al. (2013, 111) observe of proportional shortfall that “broad consensus exists in the Netherlands to use it for equity weighting.”

Conceptually, an equity weight is a factor by which some (but not all) quanta of treatment effect are multiplied (that is, differentially valued) in the name of equity. Standardly, resources are then allocated so as to *maximise in the aggregate* these differentially valued quanta of effect.<sup>21</sup> Therefore, the Dutch use of proportional shortfall as an equity weight conforms to second-type Prioritarianism *of the maximising kind*. In this respect, it resembles Bognar's allocative principle (allocate resources so as to maximise in the aggregate life years whose assigned value diminishes with age). The two allocative principles differ only in who is to be considered worse off and what mathematical function is to specify the rate at which gains to the worse off count for more.

Finally, let us briefly consider proportional shortfall in relation to the diagrammatic approach embodied in Figure 1. We have already observed that proportional shortfall does not take into account attainments, namely capacities to benefit from treatment. But as an equity weight, proportional shortfall can be used to differentially value capacities to benefit, e.g. to multiply them by a factor between zero and one. In effect, such multiplication modifies the shape of the health frontier. The point at which the modified health frontier intersects with the -1 slope can then be identified, and the corresponding allocation will maximise aggregated benefits, which now comprise differentially valued quanta of effect. This is a way of trying to maximise in the aggregate, not simply quanta of effect, but *value* more broadly conceived. Here, value has two components: the magnitude of the health gain, and how badly off is the person before enjoying the health gain. In other words, someone's capacity to benefit and someone's initial health are both treated as relevant for the purposes of just resource allocation (precisely as per Principles 7 and 8).



The Dutch use of proportional shortfall to operationalise need as an allocative principle parallels recent discussions in Britain regarding the potential use of a similar concept. Orthodox comparisons of the costs and effects of health interventions do not differentially value quanta of effect, but the British Department of Health has commissioned work to examine whether this ought to change, namely in view of the fact that society seems to place a higher value on health gains going to people who suffer a high “burden of illness” (Miners, Cairns, and Wailoo 2013). This concept can be equated with need. The Department of Health has proposed that a person’s burden of illness should be considered in terms of “the number of QALYs lost by a patient because of their condition” (Miners, Cairns, and Wailoo 2013, 10). Miners et al. (2013, 10) clarify the concept:

More formally this is described as a person’s number of projected QALYs given current treatments for their particular condition, subtracted from the expected number of QALYs given a sex-adjusted population life expectancy without the condition.

In short, someone’s burden of illness refers to how badly off someone remains compared to healthy peers, *even with the best available treatment*. (Imagine a treatment that fails to eliminate all the deleterious effects of an illness.) Insofar as the concept of burden of illness encompasses treatment effects, it contrasts with the Dutch concept of proportional shortfall. Moreover, burden of illness is quantified not as a ratio or proportion (e.g. 0.5) but as an absolute quantity of health stock (e.g. 50 QALYs).

Miners et al. propose that the concept of burden of illness can be applied, not simply to a person, but to a population that stands to benefit from a given intervention.<sup>22</sup> The population's burden of illness can be estimated (as a number of QALYs), then this number can commend an equity weight. This equity weight can then be applied like any other, e.g. as it is in the Dutch case of proportional shortfall. In effect, an equity weight multiplies the value of QALYs gained by a given intervention. In this way, quanta of effect are differentially valued according to how badly off people are. The Dutch concept of proportional shortfall and the British proposals regarding burden of illness correspond to different perspectives on the precise respect in which someone is to be considered badly off. In this way, the Dutch and British proposals differ in what constitutes need. They can each be seen as a means of operationalising need as an allocative principle, and they each represent a case of second-type Prioritarianism of the maximising kind.

## VII. CONCLUSION

A lack of clarity and agreement on the concept of equity has impeded the measurement of equity concerns, such as someone's degree of need, as part of routine economic evaluation. Hence, it has impeded the systematic integration of equity concerns into health resource allocation (Johri and Norheim 2012). For this reason, this article has clarified options for conceptualising equity or what we have called justice in resource allocation. This has included clarifying options for conceptualising need. All of the allocative principles that we examined are substantive – in two senses. First, they are not merely *formal*, in that they do attempt to specify the

substance of what makes for equals. Second, the principles are not merely *procedural*, in the sense that they do not identify the achievement of justice purely with the satisfaction of a specific (e.g. deliberative or democratic) procedure. Our illustration is intended as a spur and aid to deliberation on allocative principles. Even within the narrow confines of our illustration's particular assumptions, precisely which allocation would be just is not obvious.

Members of the public have expressed the view that society should be willing to sacrifice aggregate "health gains to ensure that the chronically ill and permanently disabled are afforded the opportunity to realise their *potential* for health, even if this is not great" (McKie and Richardson 2005, 3).<sup>23</sup> This view opposes Principle 5 (maximise aggregate health). Moreover, the view can be intended and interpreted in a number of ways – for instance, as support for:

- Principle 6 (equally exhaust each person's total capacity to benefit);
- Principle 7 (most closely equalise final health states);
- Principle 8 (equally restore health states to population norms); or
- a type of Prioritarianism (prioritise the people worst off) (Parfit 1997).

Both the empirical work and the practical resource allocation processes that engage with any such public view need to be mindful of this so as not to misinterpret it.

Of the eight allocative principles that we have illustrated, only Principle 7 (most closely equalise final health states) and Principle 8 (equally restore health states to population norms) implicitly treat as relevant attributes both a person's initial health state and a person's incremental capacity to benefit. Therefore, we have good reason to prefer Principles 7 and 8 to the other illustrated principles, insofar as we affirm (on the basis of intuition or otherwise) the moral relevance of both a person's initial health state and a person's incremental capacity to benefit.

Should resource allocation be guided by an appeal to a fair innings? If so, what constitutes a fair innings? Principle 8 (equally restore health states to population averages) represents one way of factoring into resource allocation a ‘fair innings’. Further questions then arise as to which sub-populations ought to be distinguished and what best reflects their normal health.

Quantifying health can obscure people behind their health states and encourage the equal treatment, not of people, but of quanta of health, however health is conceived (Persad, Wertheimer, and Emanuel 2009). Applying Principle 8 (equally restore health states to population norms) may be one means by which to preserve both the advantages of quantification (including precision) and the rightness of seeing people, not quanta of health, as moral equals. Another means is to consider benefits as relative to how badly off someone is. A fundamental question with respect to justice in resource allocation concerns whether benefits ought to be viewed as relative to how badly off someone is. Further questions then arise as to how much intervention effects accruing to the worse off count for more, and how this changes with improving health. We have explicated some recent efforts to answer these questions, including in Dutch and British health economics and policy circles. Each of these efforts offer different answers, and each constitutes an instance of what we have called second-type Prioritarianism of the maximising kind, since they each seek to maximise in the aggregate quanta of effect that are differentially valued in favour of the worst off. In this respect, they can each be viewed as efforts to operationalise need as an allocative principle.

## NOTES

---

<sup>1</sup> For Daniels, unsolved rationing problems include how much we ought to prioritise the worst off and how much we ought to prioritise large gains for the few over modest gains for the many. The second problem is known as “the aggregation problem” (Daniels 1993). The “rule of rescue”, namely “the imperative to rescue identifiable individuals facing avoidable death” (McKie and Richardson 2003), can be understood in the context of the aggregation problem, namely as a commitment to securing significant gains (if only for the few) over modest gains (even if for the many).

<sup>2</sup> It is important to critique the other assumptions also, but we do not do so in this article.

<sup>3</sup> We examine Principle 1 in examining Principle 5, to which it is equivalent.

<sup>4</sup> A fraction of an intervention could also correspond to the probability of winning a lottery whose prize was the whole intervention, though there could well be ethical problems with allocating treatment in this way, e.g. it might be demeaning, namely in view of the seriousness of whether or not someone is allocated the intervention.

<sup>5</sup> For argument’s sake, we put aside the formidable and much-debated conceptual and technical challenges concerning the development and use of a health metric.

<sup>6</sup> An initial health state of zero is unusual, since a zero health state usually represents death. A zero health state need not represent this, however. For instance, consider the case of blindness (zero vision) or some types of infertility (zero fertility).

<sup>7</sup> A mathematical function could be used to specify the relation.

<sup>8</sup>  $4.2 = 12/(12+5)*6$ , while  $1.8 = 5/(12+5)*6$ .

<sup>9</sup>  $0.83 = 54/65$  (A) =  $20/24$  (B).

<sup>10</sup> Again, to calculate this, we could make some assumptions regarding the mathematical function of the curve. Or we could examine the curve section as a line. In this way, by interpolating between the known data points, we can identify the point at which persons A and B share the same health state.

<sup>11</sup> One must consider what best reflects a cohort’s normal health state. For instance, it could be the average, the median, the mode, or some other aspect of the distribution of health across the cohort. Or it could be a norm derived completely without statistics. Mainly for simplicity, we focus on the average.

<sup>12</sup> These numbers represent the modelled probabilities of a live birth after one year of trying for a baby without assisted reproductive technology for women who begin at 35 and 40 years of age, respectively (Leridon 2004).

<sup>13</sup> For instance, these fundamental questions were raised in the form of “different possible conceptions of medical need” in deliberative forums with clinicians and citizens that aimed at achieving a consensus on Australian public funding for assisted reproductive technology (Hodgetts et al. 2014, 3): “On one conception, need increases with the degree of disruption to normal health”, whereas on a competing conception of medical need, “medicine ought to overcome not so much disruptions in normal health but, for instance, limits to important opportunities”, independent of whether those limits owe to what is normal or natural (Carter et al. 2013, 85).

<sup>14</sup> Williams’s emphasis.

<sup>15</sup> This contrasts with, for one, “utilitarian ageism”, which simply equates to Principle 5. Here ‘ageism’ is not intended to connote mere prejudice, namely an allocative preference that is unjustified. Instead, the word is intended to simply connote differential treatment. The question of whether the differential treatment is justified is another matter. Daniels argues that differential treatment with age need not express mere prejudice – that is, it might be justified – because people typically move through every age bracket. Insofar as we all age, “[w]e can have (1) differential treatment without (2) disparate impact that is harmful to some people and not others” (Daniels 2008, 478). By contrast, people do not typically change (let alone move through every bracket of) race, sex, gender, or sexual orientation, so differential treatment on these scores will harm some people but not others.

<sup>16</sup> It is worth provisionally tracing the concept of the whole of one’s natural life. ‘Dying of natural causes’ is a related concept. Someone could have a fair innings, accruing the requisite health stock, without living the whole of their natural life. They may be killed in an accident, for example. One’s natural life is the number of years that a particular individual will live under favourable conditions. One individual’s natural life may be 70 years, while another’s may be 110 years. In this way, the concept of the whole of one’s natural life is not reducible to the concept of a species-typical lifespan.

---

Though the species-typical lifespan is a “scaffolding fact”, or part of the background, when it comes to the concept of the whole of one’s natural life (Carter 2013).

<sup>17</sup> Parfit’s term is unfortunate insofar as egalitarians and utilitarians also prioritise: they differ precisely in the bases on which they prioritise.

<sup>18</sup> *Pace* Bognar, we could intuit that the 30-year old and the 40-year old with the same capacity to benefit (i.e. 10 additional life years) share an equal claim on the resource. This intuition could be explicated in terms of the two people being reasonably the same age, precisely as per Savulescu’s discussion of reasonably-the-same capacities to benefit. Savulescu (1998, 228, 232-233) suggests that we have as much *reason* to allocate resources to one party as to another when their capacities to benefit are reasonably similar or, in other words, “good enough” relative to one another. Parties’ capacities to benefit need not be perfectly equivalent for us to have equal reason to allocate resources to them. One party’s capacity to benefit may be greater than another’s; nevertheless, we have equal reason to allocate interventions to one person as to another if their respective capacities to benefit are reasonably similar. Only when capacities to benefit “greatly” differ do we have more reason to allocate resources to one party than to another (Savulescu 1998, 231). Savulescu’s view clearly contrasts with Principle 5, for insofar as one seeks to maximise aggregate health, one can be said to have more reason to allocate resources to the party with the greater capacity to benefit, *however small the difference*.

<sup>19</sup> Bognar presents the fourth scenario to argue against the allocative principle on which a ‘completed’ life of, say, 70 years has value beyond the simple accumulation of years: “If death before the fair innings threshold [of 70 years] is a tragedy, then the 60-year old patient should get the drug, since that way there will be only half as many tragedies.”

<sup>20</sup> Van de Wetering et al. (2013, 112) explicitly state that the linearity of the equation is simply intended as a starting point. For instance, the public may prioritise a high proportional shortfall far above a low one, and therefore commend refining the equation accordingly. Such refinement could also be a way of operationalising the rule of rescue, we may note.

<sup>21</sup> Alternatively, the cost-effectiveness threshold for publicly funding a treatment can be reduced by a factor, which comes to same: “when people stand to lose relatively more of their remaining health, a higher cost per QALY threshold is appropriate”, i.e. we are willing to pay more for these particular QALY gains, insofar as we value them more highly (van de Wetering et al. 2013, 113).

<sup>22</sup> To be precise, the burden of illness can be considered “the absolute number of QALYs that would be lost, on average, for patients for whom the new technology is being appraised given the comparator treatment within the technology appraisal, compared to patients living at full health for an average sex-adjusted life expectancy. Note there is a discussion to be had as to whether QoL [quality of life] in the absence of disease ... is considered to be perfect health (i.e. one) or is age-adjusted” (Miners, Cairns, and Wailoo 2013, 10). The discussion regarding age reflects the kind of decision to be made between Principle 7 and Principle 8.

<sup>23</sup> Their emphasis.

## REFERENCES

- Beauchamp, Tom L., and James F. Childress. 2009. *Principles of biomedical ethics*. 6th ed. Oxford: Oxford University Press.
- Bognar, G. 2015. Fair innings. *Bioethics* 29 (4): 251-61.
- Carter, D., and A. Braunack-Mayer. 2011. The appeal to nature implicit in certain restrictions on public funding for assisted reproductive technology. *Bioethics* 25 (8): 463-71.
- Carter, Drew. 2013. Morality and a scaffolding of facts. *Philosophical Investigations* 37 (1): 78-90.
- Carter, Drew, Amber M. Watt, Annette Braunack-Mayer, Adam G. Elshaug, John R. Moss, and Janet E. Hiller. 2013. Should there be a female age limit on public funding for assisted reproductive technology? Differing conceptions of justice in resource allocation. *Journal of Bioethical Enquiry* 10 (1): 79-91.
- Cookson, R., and P. Dolan. 2000. Principles of justice in health care rationing. *Journal of Medical Ethics* 26 (5): 323-9.
- Culyer, A. J. 1992. The morality of efficiency in health care – some uncomfortable implications. *Health Economics* 1 (1): 7-18.
- . 2001. Equity – some theory and its policy implications. *Journal of Medical Ethics* 27 (4): 275-83.
- . 2006. The bogus conflict between efficiency and vertical equity. *Health Economics* 15 (11): 1155-8.
- . 2007. Need: An instrumental view. In *Principles of health care ethics*, edited by R. Ashcroft, A. Dawson, H. Draper and J. McMillan. Chichester: Wiley.

- Culyer, A. J., and A. Wagstaff. 1993. Equity and equality in health and health care. *Journal of Health Economics* 12 (4): 431-57.
- Daniels, Norman. 1993. Rationing fairly: Programmatic considerations. *Bioethics* 7 (2-3): 224-233.
- . 1994. Four unsolved rationing problems: A challenge. *Hastings Center Report* 24 (4): 27-27.
- . 2000. Accountability for reasonableness. *BMJ* 321 (7272): 1300-1.
- . 2001. Justice, health, and healthcare. *American Journal of Bioethics* 1 (2): 2-16.
- . 2008. Justice between adjacent generations: Further thoughts. *Journal of Political Philosophy* 16 (4): 475-494.
- . 2009. Just health: Replies and further thoughts. *Journal of Medical Ethics* 35 (1): 36-41.
- Dolan, P., R. Shaw, A. Tsuchiya, and A. Williams. 2005. QALY maximisation and people's preferences: A methodological review of the literature. *Health Economics* 14 (2): 197-208.
- Frankfurt, Harry. 1987. Equality as a moral ideal. *Ethics* 98 (1): 21-43.
- Galatians 2:10. The Bible: King James Version.
- Hasman, Andreas, Tony Hope, and Lars Peter Østerdal. 2006. Health care need: Three interpretations. *Journal of Applied Philosophy* 23 (2): 145–156.
- Hodgetts, K., J. E. Hiller, J. M. Street, D. Carter, A. J. Braunack-Mayer, A. M. Watt, J. R. Moss, and A. G. Elshaug. 2014. Disinvestment policy and the public funding of assisted reproductive technologies: Outcomes of deliberative engagements with three key stakeholder groups. *BMC Health Services Research* 14 (1): 204.



- Hope, T., L. P. Osterdal, and A. Hasman. 2010. An inquiry into the principles of needs-based allocation of health care. *Bioethics* 24 (9): 470-80.
- Johri, M., and O. F. Norheim. 2012. Can cost-effectiveness analysis integrate concerns for equity? Systematic review. *International Journal of Technology Assessment in Health Care* 28 (2): 125-32.
- Kind, Paul, Richard Brooks, and Rosalind Rabin, eds. 2005. *EQ-5D concepts and methods: A developmental history*. The Netherlands: Springer.
- Leridon, H. 2004. Can assisted reproduction technology compensate for the natural decline in fertility with age? A model assessment. *Human Reproduction* 19 (7): 1548-53.
- McKie, John, and Jeff Richardson. 2003. The rule of rescue. *Social Science & Medicine* 56 (12): 2407-2419.
- . 2005. Neglected equity issues in cost effectiveness analysis – part 1: Severity of pre-treatment condition, realisation of potential for health, concentration and dispersion of health benefits, and age-related social preferences. Monash University, Centre for Health Economics.
- Miners, Alec, John Cairns, and Allan Wailoo. 2013. Department of Health proposals for including burden of illness into value based pricing: A description and critique. edited by NICE Decision Support Unit.
- Norheim, O. F., R. Baltussen, M. Johri, D. Chisholm, E. Nord, D. Brock, P. Carlsson, R. Cookson, N. Daniels, M. Danis, M. Fleurbaey, K. A. Johansson, L. Kipiriri, P. Littlejohns, T. Mbeeli, K. D. Rao, T. T. Edejer, and D. Wikler. 2014. Guidance on priority setting in health care (GSP-Health): The inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Effectiveness and Resource Allocation* 12: 18.

- Oxford English Dictionary. 2007. Principle, n.: Oxford University Press.
- Parfit, Derek. 1997. Equality and priority. *Ratio* 10 (3): 202-221.
- Persad, G., A. Wertheimer, and E. J. Emanuel. 2009. Principles for allocation of scarce medical interventions. *The Lancet* 373 (9661): 423-31.
- Richardson, J., K. Sinha, A. Iezzi, and A. Maxwell. 2012. Maximising health versus sharing: Measuring preferences for the allocation of the health budget. *Social Science & Medicine* 75 (8): 1351-61.
- Savulescu, J. 1998. Consequentialism, reasons, value and justice. *Bioethics* 12 (3): 212-35.
- Smith, Graham, and Corinne Wales. 2000. Citizens' juries and deliberative democracy. *Political Studies* 48 (1): 51-65.
- van de Wetering, E. J., E. A. Stolk, N. J. A. van Exel, and W. B. F. Brouwer. 2013. Balancing equity and efficiency in the dutch basic benefits package using the principle of proportional shortfall. *The European Journal of Health Economics* 14 (1): 107-115.
- Wagstaff, A. 1991. QALYs and the equity-efficiency trade-off. *Journal of Health Economics* 10 (1): 21-41.
- Williams, A. 1997. Intergenerational equity: An exploration of the 'fair innings' argument. *Health Economics* 6 (2): 117-32.