



Can we compare health states when our standards change?

Krister Bykvist¹ 

Accepted: 7 October 2023
© The Author(s) 2023

Abstract

Among health economists, who think that preferences are the correct standard of the value of health states, it is common to assume, at least implicitly, that the correct criterion of this value takes the following schematic form: H1 is a better health state than H2 iff the members of group S prefer (on average) being in H1 to being in H2. Various candidates for members of S have been proposed, including medical experts, the general public, H1-patients, H2-patients, former H1-patients, former H2-patients, or combinations of these groups. I shall argue that criteria of this form run into serious problems, if we consider cases where people's fundamental preferences change from one health state to another. I shall also show that these problems afflict hybrid views, according to which preferences are only one determinant of the value of health, the objective value of health states being another. Finally, I shall argue that a better subjectivist criterion would be something like this: H1 is a better health state than H2 iff the H1-patients would want to be in H1 more than the H2-patients would want to be in H2. According to this criterion, the value of a health state is determined by the absolute attitudes (favouring, disfavouring, neutrality) people would have towards a health state were they to be in it. Along the way, I will also present and make use of an attitude matrix framework that enables us to represent in a simple way complex information about attitudes. With this framework at hand, we can easily see the advantages and disadvantages of different subjectivist accounts of the value of health.

Keywords Comparability · Health states · Value of health

✉ Krister Bykvist
krister.bykvist@philosophy.su.se

¹ Department of Philosophy, Stockholm University and Institute for Futures Studies, Stockholm, Sweden

1 Introduction

Among health economists, it is common to measure health by eliciting preferences. For example, the utility values in the calculations of the popular QALY measure are often determined by measuring people's willingness to trade time or risks involving different health states (Burstrom et al. 2006; Nord, 1999; Weinstein & Fineberg, 1980). Respondents are asked whether they prefer remaining in a state of ill health for a period of time to being restored to perfect health but having a shorter life expectancy. Alternatively, they are asked whether they prefer remaining in a state of ill health for a period of time to choosing a medical intervention which has a certain probability of either restoring them to perfect health or killing them. Health measures, such as QALY, are then often used as a basis for decisions on allocations of health resources.

As Hausman (2015) points out, one straightforward rationale for measuring health by preferences is that preferences are seen as the correct *substantive* criterion of the *value* of health states. This rationale seems to be implicitly or explicitly assumed by many health economists (e.g., Brazier et al., 1999; Gold et al., 1996; Torrance et al., 2002). If we focus on 'unmixed' health states rather than time- or probability mixtures of health states, the preference criterion have the following general form.¹

- A health state H1 is a *better* health state than another H2 iff the members of group S *prefer* being in H1 to being in H2.
- Health states H1 and H2 are *equally good* health states iff the members of S are *indifferent* between being in H1 and being in H2.

Here the relevant notion of betterness is supposed to help us determine health improvements and deteriorations, and thereby partially determine the allocation of health resources, at least from the perspective of benevolence, respect, or autonomy. (However, I am not ruling out that other perspectives can be relevant when assessing health states, such as fairness, equality, and efficiency, but I am putting aside those perspectives in this paper.)

S is a schematic placeholder for a range of different groups, including, medical experts, the general public, H1-patients, H2-patients, former H1-patients, former H2-patients, or some mix of the aforementioned groups. Health economists are divided on the issue of which group is the relevant one. For example, Nord (1999) and Eyal (2020) argue that the relevant group comprises those who are experiencing or have experienced the health states, whereas Gold et al. (1996) argue that it is the general public.

¹ One could argue that the criterion should be broken up into two parts: (a) H1 is better than H2 iff the *health-related quality of life* of the relevant group is higher in H1 than in H2; (b) The health-related quality of life of the relevant group is higher in H1 than in H2 iff H1 is preferred to H2 by the relevant group. I will not take a stand on whether this is the right way to go. My focus is on any view that *ultimately* explains value of health in terms of preferences.

Typically, when it is said that the members of a group prefer one health state to another (or are indifferent between them) this is not to be understood as saying that *every* single member has this preference (or indifference attitude). Rather it means that the ‘aggregate attitude’ (e.g., the average attitude) of the members is one of preference (or indifference).²

The criterion is meant to be substantive and normative. So, it has to do with value-makers: it is the preferences that *make*, in the sense of *ultimately explaining why*, one health state is better than another.³ This means that the right hand side of the criterion is supposed to explain the left hand side, not the other way around. One could instead view it as an *epistemic* criterion, but I am only dealing with foundational, substantive normative issues here.

The main aims of this paper are to

- (a) show that preference criteria of this form run into serious problems, if we consider cases where the people’s preferences *change* when they move from one health state to another.
- (b) show that these problems also afflict *hybrid* or *pluralist* criteria, according to which preferences are only one determinant of the value of health, the objective value of health states being another.
- (c) argue for a better subjectivist criterion that downplays the role of preferences but still gives a prominent place to attitudes.

The discussion will be facilitated by an ‘attitudinal matrix’ framework that helps to articulate different views on the value of health states and more clearly see their implications. This framework has been used profitably in relation to the issue about changing attitudes and attitude-sensitive wellbeing.⁴ A secondary aim of this paper is to show the advantages of using it in the discussion of the value of health.

I will proceed as follows. In Sect. 2, I introduce the mattering constraint that serves as backdrop for the discussion. In Sect. 3, I introduce a simple model of health states and attitudes, inspired by Hausman. Section 4 introduces a matrix model that helps us capture evaluatively important information in a simple but telling way. In Sect. 5, I raise some important objections to the preference criterion, and, in Sect. 6, I show that my favoured account, *diagonalism*, can answer them satisfactorily. Section 7 discusses some objections to diagonalism some of which motivates some revisions of the view, and Sect. 8 ends the paper with some concluding remarks.

² There are some difficult issues surrounding the use of averaging, which I will put aside here. As Roberts and Dolan (2004) point out, an average preference measure may mask the fact that different people have radically different attitudes to the same health states, and thus, as Ubel (2000, p. 51) puts it, the measure ‘may end up truly capturing no one’s value’.

³ This is not meant to rule out a more sophisticated view, according to which *intrinsic features* of health states explain value, but preferences, in turn, explain why those features explain value. Thanks to an anonymous referee for this journal for alerting me to this possibility.

⁴ See Bykvist (2010) and (2022).

2 The mattering constraint

Before we start assessing the preference criterion, I need to say a few words about the value of health states. I will follow the defenders of the preference criterion and take for granted, what I call, 'the mattering constraint':

- The value of a health state depends at least in part on how much it *matters* to people to be in that state.⁵

This constraint can be motivated by *subjective wellbeing* (Hausman, 2015:151) but also by *respect* or *autonomy*. Importantly, since it talks about what matters to *people*, it is a 'private value' rather 'societal value', to use Hausman's terminology (Hausman, 2015: 151). The preference criterion satisfies the mattering constraint, since 'value' can be understood as comparative value between health states and 'mattering to people' can be spelled out in terms of their preferences over health states. I will show, however, that there are better ways to satisfy this constraint.

3 A simple model of health states

The objects of health evaluations are health states. A health state *type* is a combination of activity limitation types and experience types. According to the popular quality-of-life instrument EQ-5D, to which Hausman (2015) devotes considerable attention, the activity limitation types are mobility, ability to wash and dress, and the ability to do usual activities, and the experience types are pain/discomfort and anxiety/depression (EuroQol, 2023).

In contrast, a *token* health state is the exemplification of a health state type by a person at a time in a context. I follow Hausman in understanding a context as including attitudes, culture, geography, technology, occupational, and recreational structures (Hausman 2015: 67).

Now, it seems that, on any plausible view about the value of health states, the same health state type can have different values depending how it is tokened. To borrow some examples from (Broome, 2002: 95), compare having asthma in well-ventilated housing to having it in non-ventilated housing, suffering from a mental disorder in supportive community to suffering from it while living on the streets, or having severe mobility limitations but access to internet to having these limitations but no such access.

In order to sidestep the problem of assigning value to a health state type whose value varies across different times or contexts, I shall apply the subjective criterion

⁵ This rules out *pure objectivism* about the value of health, according to which certain health states can have value even though they do not matter at all to people. Whether this is a defensible view is beyond the scope of the paper. My aim here is to find a plausible view under the assumption that the mattering constraint is correct.

to health states *at a certain time and context*.⁶ More specifically, in my simple model, I will hold fixed a certain time and context for each health state type. For simplicity, I will suppress the relativization to times and contexts in what follows. Since a context is partly defined in terms of the attitudes one would have if one were in it, I will also assume:

- *Homogeneity* all Hi-persons (persons in Hi) have the same attitudes (or lack thereof) towards all relevant health states.

So, if one Hi-person has (or lack) certain attitudes towards H1 and H2 (with certain strengths), then this holds for all other Hi-persons. But people in different health states can have different attitudes towards the same health states. Homogeneity is unrealistic, but it is a modelling assumption that enables us to sidestep issues about how to aggregate different people's attitudes.

Finally, to make things simple, I will set aside preferences over mixtures of health states and only focus on preferences over unmixed health states.

In this model, the preference criterion can be formulated in this simple way:

- H1 is a better health state than H2 iff all members of group S prefer being in H1 to being in H2.

And individual preferences over health states are understood in this way:

- A person prefers being in H1 to being in H2 = df. the person prefers his H1-token to his H2-token.

I will follow the account of the polarity (valence) of attitudes set out in (Hurka, 2001: 13–14), and accept the following identities (which can be conceptual, meta-physical, or both):

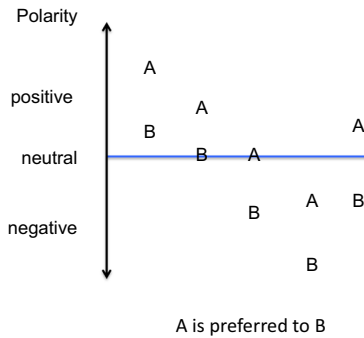
- To *favour* A = to be positively oriented towards A in your actions, emotions, feelings or evaluative responses.
- To *disfavour* A = to be negatively oriented towards A in your actions, emotions, feelings or evaluative responses.
- To be *neutral* (indifferent) towards A = to be indifferently oriented towards A in actions, emotions, feelings, or evaluative responses.

This is, of course, a very rough characterization. A more exact one requires taking a stand on exactly which attitudes are relevant. But I want to stay neutral on the exact nature of the attitudes that come with a polarity.

Finally, I will assume that the relationship between polarity and preference, seen as choice disposition, is the following (at least for rational agents with a coherent set of attitudes and preferences).

⁶ For more on the problem of assigning value to health state types, see Hausman (2015), 66–74.

Relationship between preference and polarity



Of course, this alignment can be questioned. One way to make the alignment less controversial is to focus instead on the notion of preference that is less directly connected with choice dispositions. According to this notion, to prefer A to B is to either to favour A more than one favours B, or to disfavour B less than one disfavour A, or to favour A and to disfavour B, or to favour A and be neutral towards B, or to be neutral towards A and disfavour B. Arguably, this notion has greater normative weight, since choice dispositions can come apart from what I favour most. Think of a compulsory urge to not to walk on the lines on the sidewalk, something which you do not endorse.

4 Changing attitudes and the matrix model

It is a commonplace that we have unstable health state preferences. Whether and to what degree one prefers one state to another depends on which state one is in. For example, paralysis, dialysis, colostomy, or rheumatoid arthritis is more strongly dispreferred if one is healthy than if one is in these states (Mentzel et al., 2002). Having sight or hearing is more strongly preferred if one has acquired it than if one has not. Indeed, sometimes having sight or hearing might be dispreferred by those who lack them, since being blind or deaf can be seen as an important part of one's identity.

This is not just a matter of informational change. It is true that the experiential aspect of a health state is best known when one is in it. But entering a new health state can also change one's *fundamental* preferences, for example, by altering the conception of what it means to be healthy (Mentzel et al., 2002: 2151–2152), or by replacing mobility-dependent goals with mobility-independent ones (Mentzel et al., 2002: 2151).

But one might wonder here why one's current attitudes differ from one's hypothetical ones, if you *know* that you would change your fundamental preferences? If

Fig. 1 The matrix

		Health states	
		H1	H2
Attitudes in	H1	a	b
	H2	c	d

you know that you will hate being in a certain health state, why would you now have a different attitude towards it? It is true that some of our attitudes are conditional on their own persistence, in the sense that we now, in the actual state, want something in a hypothetical state only on the condition that we would still want it to happen if we were in that state. But preferences that are associated with *ideals* can be *unconditional*, or at least only take partially into account the preferences in the considered health state. For example, the healthy and committed athlete will find the prospects of mobility limitations daunting, even though they know that they would adjust to these limitation and come to favour their life with these limitations.

In order to make clear exactly what the problems with the preference criterion is and how alternative approaches fare, it is useful to introduce a *matrix framework*.⁷ In its most simple applications, where we only have two health states to consider, the matrix looks like Fig. 1.

a is the attitude one would have towards H1 if one were in it, and d is the attitude one would have towards H2 if one were in it. These are the *intra-state attitudes*. b is the attitude one would have towards H2 if one were in H1, and c is attitude one would have towards H1 if one were in H2. These are the *cross-state attitudes*.

Polarity is represented in the following way. A *favouring* is represented with positive numbers, where a greater positive number means a stronger favouring; a *neutral* attitude with 0; a *disfavouring* with a negative number, where a greater negative number means a stronger disfavouring.

Preference is represented by one value being greater than another in the same row. For example, in H1, H1 is preferred to H2 iff $a > b$.

Indifference is represented by one value being identical to another in the same row. For example, in H1, H1 is *indifferent* to H2 iff $a = b$.

For simplicity, I will assume *full comparability* of attitude degrees across alternatives, so for any values x and y in the matrix, either $x > y$, $y > x$, or $x = y$.

As we will see later, sometimes people cannot form any attitude towards a state. In the matrix model, this is represented by the *absence* of any attitude value – a blank. So, for example, if there is a blank in the entry (H1, H2) in Fig. 1, this means that the H1-people cannot form any attitude towards H2. If there is an attitude value assigned to (H1, H1) this means that they can still form an attitude towards H1, the state there are in.

Here is a schematic example that makes use of the matrix model (Fig. 2).

⁷ See Bykvist (2010) and (2022) for similar framework applied to wellbeing and changing attitudes.

Fig. 2 Paralysis

Attitudes in	Health states	
	H1	H2
H1	2	3
H2	-1	2

Fig. 3 Blindness

Attitudes in	Health states	
	H1	H2
H1	2	1
H2	1	3

- H1 = paralysis
- H2 = no paralysis
- The H2-group can be the general public, medical experts, or former H1 patients.

From this matrix we learn that paralysed people favour their own state but favour not being paralyzed more. So, they prefer not being paralyzed to being paralyzed. Non-paralyzed people favour their own state as much as paralyzed people favour theirs, but they disfavour being paralyzed. So, they also prefer not being paralyzed to being paralyzed.

Here is another schematic example (Fig. 3).

- H1 = blind
- H2 = not blind.
- H2 can be the general public, medical experts, or former H1 patients.

From this matrix we learn that blind people favour being blind more than they favour not being blind (perhaps because they feel it is part of their identity). So, they prefer being blind to not being blind. Non-blind people favour being non-blind more than they favour being blind. So, they prefer not being blind to being blind.

5 Problems for the preference criterion

5.1 Transformative experiences

One important problem that afflicts all version of the preference criterion are transformative experiences (Paul, 2014). Sometimes in order to fully grasp a certain health state you need to be in it (or have been in it). Being in the health state provides you with the concepts that characterize the state. Entering this state would thus amount to an *epistemic* transformation. Plausible candidates for such experiences are acquiring hearing after having been completely deaf since birth, or acquiring sight after having been completely blind since birth. In these cases, the transformation involves acquiring crucial *experiential* concepts that define the state in question. But a transformative experience could also involve gaining the the concepts that describe the activity limitations in the health states under consideration.

Fig. 4 Transformative experience

Attitudes in	Health states	
	H1	H2
H1	1	
H2		2

In the matrix framework, the general problem with transformative experiences can be represented as in Fig. 4.

H2 cannot be evaluated when you are in H1, and H1 cannot be evaluated when you are in H2. So, there are no preferences to take into account. Hence, the preference criterion would generate *incomparability*—neither health state is better than the other, nor are they equally good—despite the fact that there are well-defined *intra-state attitudes*. (Note that this is a problem for all versions of the preference criterion, no matter whose preferences you consider, since we can assume that whatever health state, H3, is added, if one were in that state one would still not be able form a preference between H1 and H2.)

Can a hybrid view rescue the preference criterion? According to a hybrid view, how one health state compares to another depends not only on one’s preferences about them but also on their objective values. Such a view will not help, however, if value is seen as a function of both preferences and objective value. If there are no preferences, the function is not well-defined.

Of course, a more *pluralist* view could help, since it could say that objective values call the shot if there are preferential gaps. So, we would get an value ordering even in the absence of preferences. But absolute intra-state attitudes are still excluded, even when they are well-defined. Indeed, the fact that one were to favour H2 more in H2 than one would favour H1 in H1 speaks in favour of H2 over H1. But this cannot be taken into account, if the only attitudes that can make a difference are preferences between states. Indeed, not taking these attitudes account seems to run counter to the mattering constraint, according to which the value of health states depend in part on what matter to people.

5.2 Mismatch between preferences and absolute attitudes

Another problem for the preference criterion is the possible *mismatch* between preferences and absolute intra-state attitudes. Here is one version of this problem (Fig. 5).

The preference criterion would have to say H1 and H2 are *equally* good, since no matter which state one were in one would be indifferent between the two states. But

Fig. 5 Shared comparative indifference

Attitudes in	Health states	
	H1	H2
H1	2	2
H2	-2	-2

Fig. 6 Shared preference

Attitudes in	Health states	
	H1	H2
H1	2	3
H2	-2	-1

one would *favour* H1, if one were in H1, but *disfavour* H2, if one were in H2, which clearly speaks in favour of H1 over H2.

Here is another version of the problem of the mismatch between preferences and absolute intra-state attitudes (Fig. 6). The preference criterion would have to say H2 is better than H1, since no matter which state one were in, one would prefer H2 to H1. But one would *favour* H1, if one were in H1, but *disfavour* H2, if one were in H2, which clearly speaks in favour of H1 over H2.

6 Diagonalism

Diagonalism is a subjective criterion of the value of health states that avoids all objections raised against the preference criterion.⁸ According to this view, the absolute intra-state attitudes call the shots. Roughly, H1 is a better health state than H2 iff the H1-persons *would* want to be in H1 more than the H2-persons *would* want to be in H2. A more precise formulation can be given if we turn to this matrix (Fig. 7).

- H1 is better than H2 iff $a > d$.⁹
- H1 is equally as good as H2 iff $a = d$.
- H1 is good iff $a > 0$; bad iff $a < 0$; neutral iff $a = 0$.
- H2 is good iff $d > 0$; bad iff $d < 0$; neutral iff $d = 0$.

Diagonalism can easily deal with the two main objections to the preference criterion. First, consider preferential gaps (Fig. 8).

H2 is better than H1, since the favouring of H2 in H2 is stronger than the favouring of H1 in H1. It does not matter that neither the H1-people nor the H2-people can form a preference over these states.

Diagonalism also easily deals with the mismatch between preferences and absolute intra-state attitudes (Fig. 9). H2 is worse than H1, since H2 would be disfavoured in H2 and H1 would be favoured in H1. H2 is worse than H1, since one would disfavour H2, if one were in H2, and one would favour H1, if one were in H1 (Fig. 10).

Even though my main target in this paper is the preference criterion, one might wonder why we should not count all absolute attitudes towards a certain health state, both the intra-state ones and the inter-state ones. This case, which we have already introduced, shows why we cannot count them *equally* (Fig. 11):

⁸ See Bykvist (2010) and (2022) for objections to the preference criterion of wellbeing.

⁹ I assume, as is customary, that 'worse' is the converse of 'better'. Hence there is no need for a separate clause for worseness.

Fig. 7 Diagonalism

		Health states	
		H1	H2
Attitudes in	H1	a	b
	H2	c	d

Fig. 8 Diagonalism and trans-formative experience

		Health states	
		H1	H2
Attitudes in	H1	1	
	H2		2

Fig. 9 Diagonalism and shared comparative indifference

		Health states	
		H1	H2
Attitudes in	H1	2	2
	H2	-2	-2

The absolute attitude towards H1 balance each other out, if the number of H1-persons is the same as the number of H2-persons, and we can make cardinal comparisons of attitudes. The same holds for the absolute attitudes towards H2. So, on this view, H1 and H2 are equally good, which is the wrong result.

This case does not refute the view that inter-state attitudes have *some* weight but *less* than intra-state attitudes. Here is a schematic case that makes for troubles for this more tempered view, if it still assumes that inter-attitude can outweigh intra-state attitudes (Fig. 12).

For some values of -n and m, H2 will have to be seen as better than H1, even though H1 would be strongly favoured (we assume) in H1 and H2 would be strongly disfavoured in H2. This is still counterintuitive, at least if we assume that the H1-people and H2-people are equally well-informed about the two states.

A more general problem is that any view that takes into account *all* inter-state attitudes and give them some weight, minimally as tie-breakers, will have to accept that whether one state is better than another depends on which *other* states are seen as alternatives. So, it is not enough to look at the attitudes in H1 and H2 in order to decide which state is better, not even if the H1-people and H2-people are in *complete agreement* about how they feel about these two states. The attitudes in a third state, H3, can make a difference, minimally as a tie-breaker, if they are directed at H1 or H2. This is counterintuitive.

Furthermore, how do we decide which health states are relevant alternatives? All logically possible states? No, that is not feasible. But how do we delineate the relevant set? Relative to some agent who can realize these states? But exactly which agent? Note also that expanding or shrinking the set can make a difference to how we should value the two states that are retained. So, the ranking of two states is potentially unstable across time, which makes for trouble when implementing plans

Fig. 10 Diagonalism and shared preference

Attitudes in	Health states	
	H1	H2
H1	2	3
H2	-2	-1

Fig. 11 Counting all absolute attitudes equally

Attitudes in	Health states	
	H1	H2
H1	2	2
H2	-2	-2

Fig. 12 Counting inter-state absolute attitudes less

Attitudes in	Health states	
	H1	H2
H1	10	m
H2	-n	-10

for allocation of health resources. Diagonalism will avoid all these problems, since, on this view, the value of a health state does not depend on the set of alternative states.¹⁰

In order to avoid the dependence on the whole option set, one could instead say that whether a state is better than another just depends on those attitudes, in these two states, that are directed at these two states. However, as I will show in the next section, this pairwise comparison approach will lead to cyclical value orderings, even for the minimal view that only treats inter-state attitudes as tie-breakers.

7 Objections

7.1 Preferences as tie-breakers?

One could object to diagonalism, because it throws away all preference information (and all information about inter-state attitudes) (Fig. 13).

Diagonalism would have to say that H1 is equally as good as H2, even though in H1 one would be in a state one prefers to be in, whereas in H2 one would be in a state one disprefers to be in. So, isn't it better to at least let the preferences in two states work as *tie-breakers* in those cases where the intra-state attitudes in those states are tied? No, for this leads to cyclical health values (Fig. 14).

H1 is better than H2, since, even though the absolute attitudes in H1 and H2 are tied—H1 is favoured to degree 2 in H1 and H2 is also favoured to this degree in H2—both the H1-people and the H2-people prefer H1 to H2—the favouring degrees

¹⁰ For more on this kind of dependence, see Bykvist (2010).

Fig. 13 Tie-breaking

Attitudes in	Health states	
	H1	H2
H1	2	1
H2	3	2

Fig. 14 Cyclical betterness

Attitudes in	Health states		
	H1	H2	H3
H1	2	1	3
H2	3	2	1
H3	1	3	2

are 2 and 1, and 3 and 2, respectively. The same reasoning applies to H2 and H3, and H3 and H1, respectively. So, we get a cycle: H1 is better H2; H2 is better than H3; but H3 is better than H1.

Though it is a contested issue whether circular betterness is conceptually impossible, it is definitely not an attractive feature of health state value.¹¹ It sometimes makes it impossible to avoid being in a suboptimal health state. For example, in the case above, whichever health state people are in there is better health state they could be in. Circular betterness thus makes it difficult to use health value as a guide to the allocation of health resources, since there is no optimal health state to aim at.

7.2 Adaption problems

If we let the absolute H_i -attitudes decide the value of H_i , we have to face obvious adaption problems. Sometimes adaption to a state can be innocent, or even beneficial if one gains crucial information about the state that it is hard to get without having experienced the state, for example, the specific qualitative feel that partly defines the state.

But sometimes adaption seems problematic. Adaption-disfavouring factors include factual mistakes, such as cognitive denial, failure to recognize full health (Mentzel et al., 2002: 2151), and mistaken goals, such as lowered expectations, the extreme version being ‘happy slave’ and ‘sour grapes’ (Mentzel et al., 2002: 2152).

To deal with factual mistakes, one can correct for false information by taking into account absolute attitudes of former patients, as (Eyal, 2020) suggests, and perhaps even parts of the general public.¹² This does not mean that we abandon diagonalism, since these attitudes can be seen as an *approximation* of the *hypothetical* attitudes of the patients – the ones she would have if she were well-informed—when there is enough agreement between the fundamental preferences of patients and former patients (or parts of the general public).

¹¹ For arguments for the possibility of circular betterness, see Temkin (2011). For scepticism about this possibility, see Broome (2004: ch. 4).

¹² For a critical discussion of Eyal’s proposal, see Mosquera (2021).

In order to solve the problem with mistaken goals, a more radical revision is required. We need to go hybrid and let the objective value of a state in part determine its value. It is important that objective value only *partly* determines value because otherwise we would violate the mattering constraint on which my discussion is premised.

- *Hybrid view*: the value of a health state H_i depends on the *match* between the attitudes in H_i towards H_i and the objective value of H_i .

On this view, a whole-hearted strong endorsement of a health state that is objectively very bad does not make it a good health state, since this is a complete mismatch of attitudes and objective value. Similarly, a whole-hearted strong disapproval of an objectively very good one does not make it bad, since this is also a complete mismatch. Exactly how the match between attitudes and objective values determines the value of health states is a difficult issue and there is room for many different accounts. But it seems at least clear that any plausible account must accept that the greater objective value a favoured health state has, the higher health value it will have.

It also seems plausible to assume that the more an objectively good or neutral health state is favoured, the better it is, and the more an objectively bad or neutral state is disfavoured the worse it is. Of course, this does not exhaust all combinations of attitudes and values. It is beyond the scope of this paper to work this out in detail, but a more complete account could borrow some ideas from Kagan (2021), who discusses similar issues in relation to wellbeing.

It should be noted that the adaption problem is a problem for the preference criterion as well, since preferences can also be adapted in problematic ways. So, we can hardly say that the preference criteria has the upper hand here. Indeed, here the preference criteria and diagonalism are in the same boat.

7.3 Incomparability: attitude-less states

While the preference criterion has problems assessing health states that involve a transformative experience, diagonalism has problems with assessing states that a person would not be able to form any absolute attitudes towards, if she were in them. Think of very severe dementia or cognitive impairments, or being in a coma. Being in such a state might prevent one from forming an attitude towards one's state. But we don't want to say that the state lacks value and is neither good, bad, nor neutral, since being mentally and physically healthy seems better than being in a coma, or suffering from severe cognitive impairments. But diagonalism, as it stands, requires that a state has value only if the persons in it has some attitude towards it, favouring in case of goodness, disfavouring in the case of badness, and neutral attitude in the case of neutral value.

This issue motivates a slight revision of diagonalism. Instead of insisting that a state has neutral value iff persons in it would have a neutral attitude towards it, we should say that it has neutral value iff *either* people in it would have a neutral attitude towards it *or* they would *lack* any attitudes towards it. This is not an ad hoc

move, for note that standard wellbeing theories, such that hedonism, would say something analogous about neutral wellbeing. According to standard hedonism, a person has neutral wellbeing iff either the person experiences hedonic indifference, or she lacks any positive, negative, or indifferent experiences. It is important that when we say that a person lacks an attitude we are saying that the person *exists* and has no attitudes. Non-existence is not a state that can have health value, since it is not a state one can be in.

One should also remember that in severe cases of dementia and cognitive dysfunction it is often possible to have likes and dislikes towards the *experiential* aspects of the state, for example, a desire for warmth and a gentle touch, and a dislike for pain. This is enough to define the value of the health state, at least in a rough way.¹³ Even in some cases of coma the grounds for the dispositional attitudes are still there, but because of the coma they cannot be manifested. It remains true that being in coma without any dispositional attitudes at all is neutral for the patient. But that seems reasonable, since it implies that it is better to prevent suffering, a bad state, than being in a coma. Furthermore, there is no extra health value to be gained by keeping the comatose patient alive (if there is no prospect of her waking up and forming new attitudes). Of course, this does not rule out that the patient herself or her family and friends may have a legitimate *non-health* related interests in keeping her alive.

It is important to note that proponents of the preference criterion cannot avail themselves to a similar solution to their incomparability problem (discusses in Sect. 5.1). The problem for them was that two states have to be seen as incomparable in value because the relevant group of people could not form a comparative attitude towards these states, despite the fact that the people in each state it would have attitudes towards it. If the criterion was revised so that two states are neutral whenever the relevant group has no preference towards them, it would mistakenly say that H1 and H2 in Fig. 4 and Fig. 8 are equally good, when it is clear that H2 is better than H1, since one would favour H2 more, if one were in it, than one would favour H1, if one were in it.¹⁴

8 Concluding remarks

The preference criterion is flawed, and it cannot be saved by a hybrid account. Diagonalism is a better subjectivist criterion. But the best version needs to go hybrid. However, much more needs to be said about the details of such a view. Once we have seen those details we are in a position to decide whether the hybrid version of diagonalism is not just preferable to the preference criterion, which is what I claim

¹³ I am open to the possibility that the value of a health state can be indeterminate to some degree. Perhaps we can only say that is good (bad), or good (bad) within a certain range of positive (negative) values.

¹⁴ Thanks to an anonymous referee for raising this issue.

here in this paper, but also preferable to non-subjective accounts of the value of health.

To those who have some doubts about my conclusions, I hope that they at least appreciate the matrix framework that I have presented, since it provides a very simple model of changing health attitudes that succinctly captures relevant information about those attitudes, and partly because of its simplicity, makes it much easier to compare the advantages and disadvantages of different attitude-sensitive theories about the value of health.

Acknowledgements For incisive and very helpful comments, I would like to thank the two anonymous reviewers of this journal and the audience at the workshop on incommensurability and population-level bioethics, at the Centre for Population-Level Bioethics, Rutgers University, May 17, 2022. I am especially grateful for comments from Nir Eyal, Dan Hausman, Anders Herlitz, Frances Kamm, and Wlodek Rabinowicz.

Authors' contribution Not applicable, since it is single-authored.

Funding Open access funding provided by Stockholm University. This work was supported by Riksbankens Jubileumsfond, M170372:1

Data availability Not applicable, since no data was used. Only references to published work.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

Ethics approval No ethics approval needed, since no sensitive data was used, only references to published academic work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Brazier, J., Deverill, M., Green, C., Harper, R., & Booth, A. (1999). A review of the use of health status measures in economic evaluations. *Health Technology Assessment*, 3, 1–164.
- Broome, J. (2002). *Measuring the burden of disease by aggregating well-being* In C. Murray, C. J. Solomon, C. Mathers, & A. Lopez, (Eds.), *Summary measures of population health: Concepts, ethics, measurement, and applications* (pp. 91–113). World Health Organization.
- Broome, J. (2004). *Weighing lives*. Oxford University Press.
- Burstrom, K., Johannesson, M., Diderichsen, F., et al. (2006). A comparison of individual and social time trade-off values for health states in the general population. *Health Policy*, 76(3), 359–370.
- Bykvist, K. (2010). Can unstable preferences provide a stable standard of well-being? *Economics and Philosophy*, 26(1), 1–26.

- Bykvist, K. (2022). Wellbeing and changing attitudes across time. *Ethical Theory and Moral Practice*. <https://doi.org/10.1007/s10677-022-10311-x>
- EuroQol. (2023). <https://euroqol.org/>. Accessed 24 April 2024
- Eyal, N. (2020). Measuring health-state utility via cured patients. In I. Glenn Cohen, C. Shachar, A. Silvers, & M. A. Stein (Eds.), *Disability health, law, and bioethics* (pp. 274–277). Palgrave.
- Gold, M., Patrick, D., Torrance, G., Fryback, D., Hadorn, D., Kamlet, M., et al. (1996). Identifying and valuing outcomes. In M. Gold, J. Siegel, L. Russell, & M. Weinstein (Eds.), *Cost-effectiveness in health and medicine: Report to the U.S. public health service, panel on cost-effectiveness in health and medicine* (pp. 82–134). Oxford University Press.
- Hausman, D. (2015). *Valuing Health. Well-Being, Freedom, and Suffering*. Oxford University Press.
- Hurka, T. (2001). *Virtue, vice, and value*. Oxford University Press.
- Kagan, S. (2021). What is the opposite of wellbeing? In J. McMahan, T. Campbell, J. Goodrich, & K. Ramakrishnan (Eds.), *Principles and persons the legacy of Derek Parfit*. Oxford University Press.
- Mentzel, P., Dolan, P., Richardson, J., & Olsen, J. A. (2002). The role of adaption to disability in health state valuation: A preliminary normative analysis. *Social Science and Medicine*, 55, 2149–2158.
- Mosquera, J. (2021). The bias of adapted patients in practice. *Journal of Law and the Biosciences*, 8(2), 1–10.
- Nord, E. (1999). *Cost-value analysis in health care: Making sense of out of QALYs*. Cambridge University Press.
- Paul, L. (2014). *Transformative experience*. Oxford University Press.
- Roberts, J., & Dolan, P. (2004). To what extent do people prefer health states with higher value? A note in evidence of from the EQ-5D valuation set. *Health Economics*, 13, 733–737.
- Murray, C, Solomon, C. J., Mathers, C., & Lopez, A. (Eds.). (2002). Summary measures of population health: Concepts, ethics, measurement, and applications. World Health Organization.
- Temkin, L. (2011). *Rethinking the good*. Oxford University Press.
- Torrance, G., Capretta, J. C., Miller, T. P., & Moffit, B. (2002). Health utility estimation. *Expert Review of Pharmacoeconomics Outcomes Research*, 2, 99–108.
- Ubel, P. (2000). *Pricing life: Why it's time for health care rationing*. MIT Press.
- Weinstein, M. C., & Fineberg, H. V. (1980). 'Clinical decision analysis.' W. B. Sanders.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.