

IS NICE TOO SEVERE WITH SEVERITY? Exploring how well NICE's severity modifier aligns with UK public preferences

OHE Helen Hayes, PhD Priscila Radu David Mott, PhD Chris Skedgel, PhD cskedgel@ohe.org

ABPI Orban Holdgate Victoria Jordan

BACKGROUND & AIMS NICE, like other HTA bodies, has implemented a severity modifier that assigns greater value to health gains for patients with greater **absolute (AS)** or **proportional (PS)** health shortfalls.¹

Under NICE's current criteria, based on the more severe of absolute or proportional QALY shortfall, patients with a particular disease qualify for a value multiplier of 1.2 if they are expected to lose between 85% and 95% of their (discounted) expected lifetime QALYs or between 12 and 18 (discounted) QALYs relative to an individual without the disease.¹ Patients are eligible for a multiplier of 1.7 if they are expected to lose more than 95% of their (discounted) expected lifetime QALYs or more than 18 (discounted) QALYs relative to an individual without the disease.¹

NICE is committed to accounting for societal values in its methods, but to date it has provided little empirical justification for its severity thresholds or value multipliers.²

Our aim, therefore, was to understand how well NICE's current severity modifier aligns with UK societal preferences over the relative value of health gains to patients with different levels of absolute (AS) and proportional (PS) health shortfall:

METHODS We used a **Person Trade-Off (PTO)** approach (see example task to the right) to understand the value of health gains to patients with a greater versus lesser future health in an agegender representative sample of the England & Wales general population (N=997; complete case analysis N=990). We also elicited their views on severity thresholds.

In each task, respondents were initially presented with two hypothetical patient groups of 100 patients each. Respondents were told the groups were identical, with the exception of their future health shortfall. They were asked to imagine themselves as a health system decision-maker with enough resources to only treat one of the groups, and asked which group they would treat. Following this choice, the number of patients in the *non-prioritised group* was increased according to a fixed algorithm and the question was repeated. This continued until respondents reached an indifference point or they reached a maximum of 1000 patients in the non-preferred group. We calculated respondent-level severity weights as the ratio of patients in the more severe group to patients in the less severe group and summarised the mean and median of these respondent weights.



Based on feedback from 'think-aloud' piloting, we presented health shortfall in terms of *future health units lost* and avoided mentioning quality-adjusted life years. To help respondents understand future health lost, we explained that shortfalls could result from lost *quality-of*life, lost length-of-life, or some combination of the two sources. We did not specify the source of the shortfall in the PTO tasks.

AS = Remaining QALY expectation in absence of disease - Remaining QALY expectation with disease

Disease related QALY loss (AS) PS = -Remaining QALY expectation in absence of disease

FIGURE 1: EXAMPLE PTO TASK



from today

RESULTS

FIG 2: DERIVED SEVERITY WEIGHTS VERSUS NICE'S CURRENT SEVERITY MODIFIER Relative to 20% proportional shortfall reference group



Based on Figure 2, we find that societal concern for severity appears to begins at lower shortfalls than those currently used by NICE, and that the public assigns greater relative value to health gains to more severe patient groups than NICE's current weights. We find a fairly rapid increase in the relative value of treating more severe patient groups, even at relatively moderate levels of shortfall, and that these weights plateau around 65% PS.

Given the open-ended nature of PTO responses, we are not surprised our mean estimates are substantially higher than the medians. We suggest that median values are likely to be more representative of societal preferences, but even with these more conservative estimates are considerably higher than the NICE weights over most of the shortfall range (with the exception of the median value for 95% PS).

This interpretation is supported by complementary evidence from the elicitation of thresholds illustrated in Figure 3.

Respondents indicated where they thought 'severe' and 'very severe' health shortfalls begin on proportional and absolute shortfall scales.

On the proportional shortfall (PS) scale, the mean threshold for 'severe' was 48% PS and for the mean for 'very severe' was 64%.

FIGURE 3: DISTRIBUTION OF ELICITED PS & AS THRESHOLDS FOR 'SEVERE' AND 'VERY SEVERE'



Respondents saw this task before the PTO tasks but the ~65% PS threshold aligns closely with the value profile in Figure 2.

CONCLUSION

These results suggest that the UK public's preference for prioritising health gains in more severe health states begins at a substantially lower shortfall threshold, and that they assigns greater relative value to health gains at almost every level of severity, than NICE's current severity modifier. Our median results suggest a concave value function, with the relative value of health gains peaking between 65% - 90% PS rather than at the most extreme shortfall. Together, these results suggest that NICE's current modifier may undervalue severity in terms of eligibility and value relative to public preferences.

¹ NICE (2022) - NICE health technology evaluations: the manual. https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741 ² Hausman, D.M. (2024) Problems with NICE's severity weights. Social Science & Medicine, 348, p.116833. 10.1016/j.socscimed.2024.116833.



