Comparative Analysis of Pan-European Severity Assessment Methods and Their Uncertainty in Cost-Effectiveness Analysis

EE135

Mclean A,[1] van Hest N[1]

[1]Costello Medical, Bristol, UK

Objective

To assess the variability and probabilistic uncertainty in approaches taken by European health technology assessment (HTA) bodies to quantitively evaluate disease severity.

Background

- Disease severity is often considered by HTA bodies as part of reimbursement decisions, either qualitatively within submission dossiers or quantitively within cost-effectiveness models (CEMs). Quantitative methods typically use absolute or proportional shortfall (AS or PS) as measures of disease severity.
- Greater weight is often placed on health gains in more severe diseases in line with societal preference to prioritise health gains for those with a higher burden of illness or disability.
- This research compares the impact of the quantitative approaches currently in use by three European HTA bodies, as well as the impact of calculating severity within probabilistic analysis.

Methods

- Severity weighting methods adopted by the Norwegian Medical Products Agency (NoMA, Norway), Zorginstituut Nederland (ZIN, Netherlands) and the National Institute for Health and Care Excellence (NICE, UK), were applied in an oncology CEM for three subgroups with varying prognosis.
- **Figure 1** shows the AS and/or PS categorisations adopted by each agency and the corresponding modifier considered by the HTA body when evaluating cost-effectiveness.
- Uncertainty in shortfall estimates was explored, with deterministic and probabilistic weightings generated for each methodology:
  - General population utility and probability of death per cycle, both disaggregated by age, were informed by relevant UK databases for all HTA settings and sampled within a Monte-Carlo simulation (1,000 iterations). Each utility and mortality input was varied using a beta distribution with a standard error equivalent to 20% of the mean. The logical ordering of inputs was retained by using the same random number within each simulation for each input category.
  - The number of simulations which satisfy the conditions needed for a specific multiplier were recorded, presented as probabilities within **Figure 2**.

Results

- While each subgroup qualified for a unique deterministic NICE severity weighting (×1, ×1.2 and ×1.7), all subgroups qualified for the same modifier for ZIN (×4) and NoMA (×1.8), **Figure 2A–2C**.
  - ZIN applies higher weightings compared to NICE or NoMA for all three subgroups with the highest possible modifier applied to all three subgroups. ZIN only considers PS.
  - NoMA applies weightings based on AS only and similar to ZIN, applies the same weighting (×1.8) to all 3 subgroups, however dissimilar to ZIN, ×1.8 is not the highest possible modifier (3rd on a range from ×1 – ×3).
- The most probable modifier from the probabilistic analyses does not always align with the deterministic value. For example, 51.5% of simulations for the subgroup with poorest prognosis (subgroup 3) qualified for a NoMA weighting of 2.2, despite a deterministic weighting of 1.8 (**Figure 2C**). The likelihood of adjacent modifiers being considered can be within a few percentiles of each other, e.g., probabilistic results indicated the subgroup with best prognosis (subgroup 1) qualified for a NICE weighting of 1.2 in 49.2% of simulations, despite a deterministic weighting of 1 (**Figure 2A**).

Conclusion

There is variation between European HTA bodies on how disease severity is considered. NICE prioritises health gains in severe disease areas to a lesser extent than NoMA and ZIN, which apply higher modifiers across a wider range of disease severity.

When uncertainty in general population mortality and utility are captured, the probability of qualifying for adjacent modifiers can be within a few percentiles. This demonstrates that evaluating severity modifiers deterministically fails to capture underlying variation in modelled data and a potential need to evaluate severity modifiers via probabilistic analysis within CEMs.

FIGURE 1

Pan-European severity modifiers

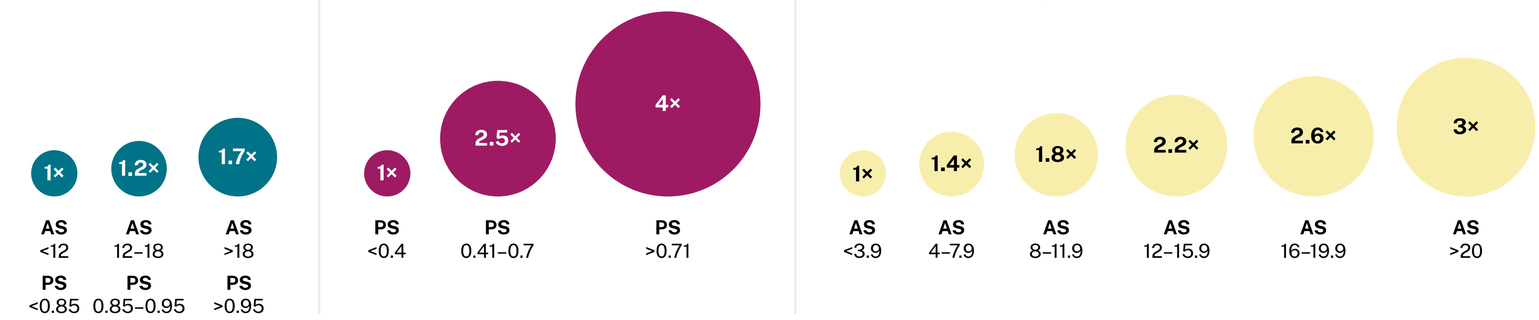

FIGURE 2

Deterministic and probabilistic severity weightings according to subgroups by prognosis

For the UK, when separate modifiers are calculated using AS and PS methods, the highest modifier is chosen in line with NICE guidance.

Abbreviations: AS: absolute shortfall; **CEM:** cost-effectiveness model; **HTA:** health technology assessment; **NICE:** The National Institute for Health and Care Excellence; **NOMA:** Norwegian Medical Products Agency; **PS:** proportional shortfall; **ZIN:** Zorginstituut Nederland.

Acknowledgements: The authors thank Ashleigh Farthing, Costello Medical, for graphic design assistance.

Presented at ISPOR Europe 2025 | Glasgow, UK | 9–12 November