HTA371

Why TSD26 Matters: How Accurate was Clinical Expert Opinion in Estimating Long-Term Outcomes?

Beale R,[1] Freeman H,[1] Porteous A[1]

[1]Costello Medical, London, UK

Objective

To investigate the accuracy of clinical expert estimates of long-term survival outcomes prior to the publication of TSD26, and to compare the reliability of different expert elicitation methods used.

Background

- The recent National Institute for Health and Care Excellence (NICE) Decision Support Unit (DSU) Technical Support Document (TSD) 26 details robust methods for eliciting long-term survival outcomes from clinical experts.[1] Given the significant impact survival outcomes have on cost-effectiveness, it is vital predictions are accurate.

Methods

- NICE appraisals published between January 2020 and December 2023 in the four most common cancers (breast, lung, prostate and colorectal) were reviewed for use of clinical expert opinion to inform long-term survival extrapolations.
- For each identified appraisal, eligibility for the analysis was assessed by capturing the method used for expert elicitation, the availability of experts' estimates (i.e. redacted or unredacted) and the availability of longer-term data relative to the earlier trial data-cut used to inform expert validation.
- For appraisals where clinical expert estimates and subsequent longer-term trial data were available, the estimated progression free survival (PFS) or overall survival (OS) estimates by clinical experts were extracted, either directly from the committee papers or through digitizing validated survival curves. Longer-term trial data were then extracted from the literature and compared to clinical expert estimates.

Results

- 49 appraisals were identified across breast (n=13), prostate (n=8), lung (n=23) and colorectal (n=5) cancer, of which 44 included expert opinion to inform extrapolation of survival outcomes. Expert opinion was elicited through telephone conferences (n=26), advisory boards (n=17) and surveys (n=1). No appraisals were identified that used formal methods for structured elicitation in line with the recommendations from TSD26.
- Nine appraisals used expert opinion, had undredacted expert estimates, and long-term data were available from the clinical trial used to inform expert validation (**Figure 1**).
- Across these nine appraisals, 29 survival timepoints were evaluated across PFS and OS (**Table 1**).
- Overall, clinical expert estimates were misaligned (defined as >5% difference from trial data) with realised survival at 13 of 29 timepoints, with underestimation at 11 timepoints and overestimation at two. There was no association between the time elapsed between data cuts and the accuracy of estimates. For the remaining 16 timepoints, clinical expert survival estimates were within 5% of long-term trial data.
- **Figure 2** provides an overview of the accuracy of the clinical expert estimates by method of elicitation.
- Validation of extrapolation curves was used in 19 of the 29 timepoints identified and expert predictions were elicited in the remaining 10 timepoints. In general, expert predictions were less accurate at estimating future survival compared to experts validating extrapolated curves.

Conclusion

Our research demonstrates clinical expert estimates were frequently misaligned with realised long-term survival, with systematic underestimation in the vast majority of misaligned cases. This suggests inherent limitations in current expert elicitation methods and potential conservative bias that may impact cost-effectiveness evaluations and reimbursement decisions. Given the large impact survival outcomes have on cost-effectiveness outcomes, more structured approaches as outlined in TSD26 may improve elicitation methods and potentially enhance prediction accuracy.

FIGURE 1

Summary of appraisals identified

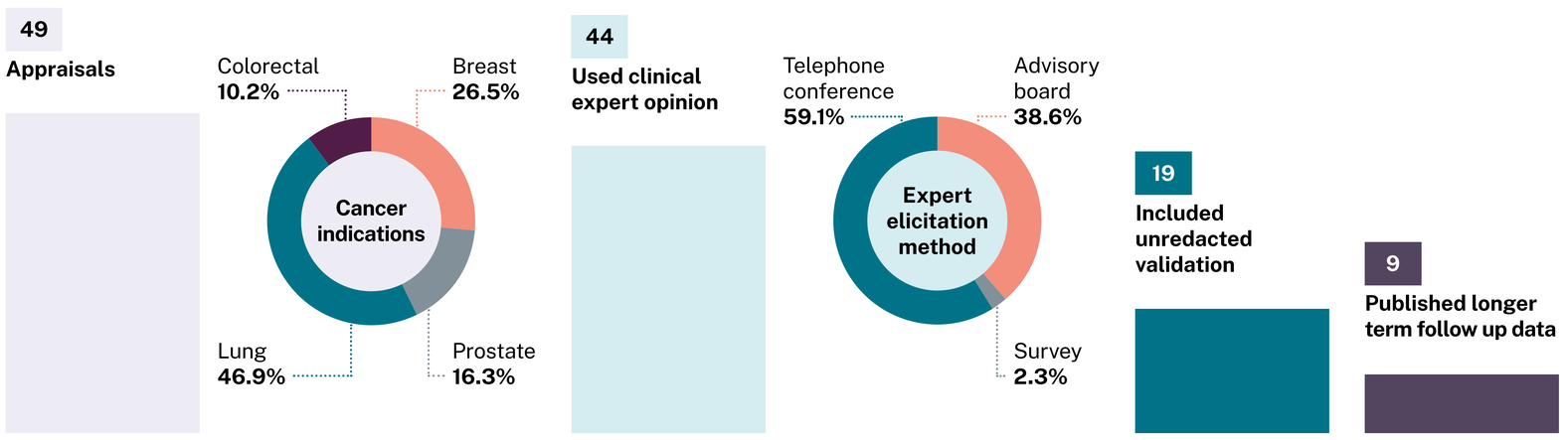


TABLE 1

Comparisons between clinical expert survival estimates and long-term trial data

TA	Trial	Intervention/comparator	Time between datacuts (months)	Endpoint	Difference in % survival between clinician estimates and realised trial data at available timepoints (months) 30	36	42	48	54	60
TA862	DESTINY-Breast03	Trastuzumab Emtansine (T-DM1)	30.9	OS						−9.3% to −19.3%
TA786	HER2CLIMB	Tucatinib plus trastuzumab and capecitabine	17.3	OS		+6.9%		+3.6%		
				PFS		+0.8%		−4.2%		
TA687	MONALEESA-3	Ribociclib plus fulvestrant	17.0	OS				+0.3%	−2.6%	−6.8%
		Placebo plus fulvestrant		OS				−3.1%	−6.3%	−4.4%
TA639	IMpassion130	Atezolizumab with nab-paclitaxel	15.5	OS	−6.9%	−6.2%		−16.7%		
TA660	ARAMIS	Darolutamide plus ADT	14.5	OS				+4.3%	+1.5%	+9.6%
		ADT		OS				−1.0%	−3.8%	−11.1%
TA911	LIBRETTO-001 (SAS1)	Selpercatinib	19.1	OS		−5.6%				
				PFS		−2.1%				−2.3% to −7.3%
TA823	IMpower010	Best supportive care	37.1	OS						−10.3%
TA670	ALTA-1L	Brigatinib	18.3	OS		−4.2%	−5.8%	−7.8%		
TA709	KEYNOTE-177	Pembrolizumab	41.2	PFS						−4.0% to +16.0%
		Chemotherapy		PFS						−2.6% to +2.4%

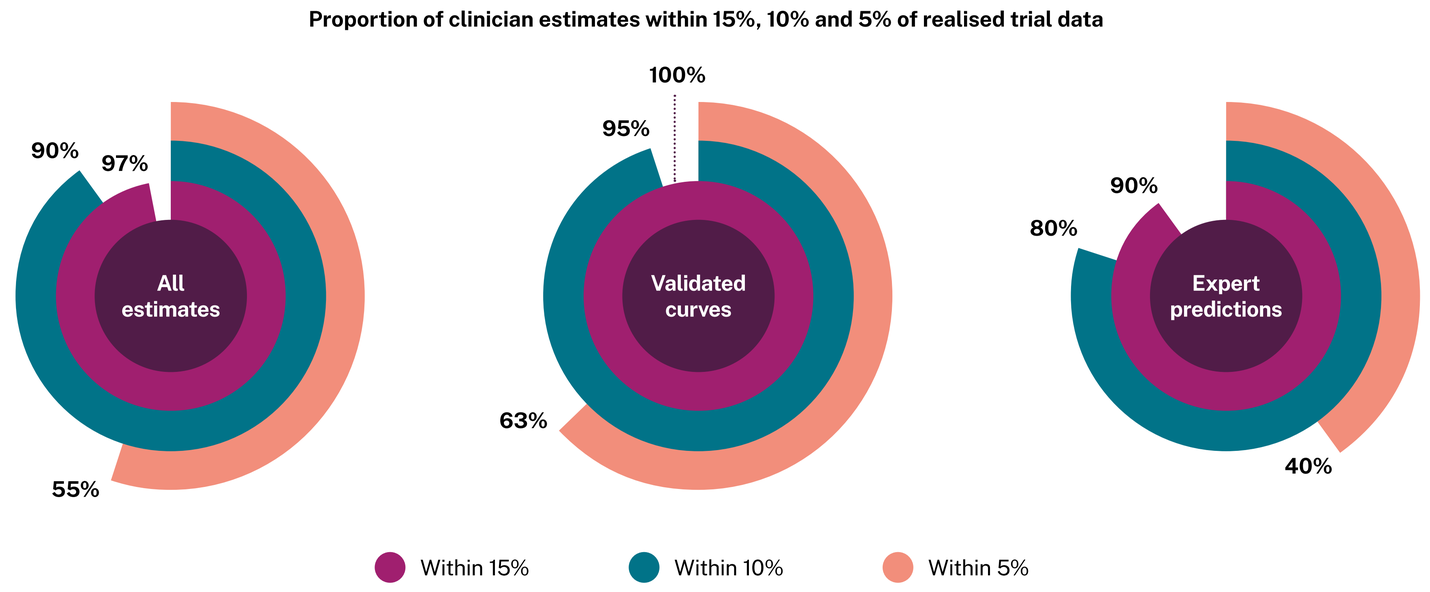
Clinical expert estimate was within 5% of trial data

Clinical expert estimate was overpredicted by >5% compared to trial data

Clinical expert estimate was underpredicted by >5% compared to trial data

FIGURE 2

Comparison of clinical expert elicitation methods



Validated curves refers to methodologies where clinical experts have chosen or agreed with a presented extrapolated curve beyond the KM data from the trial.
Expert predictions refers to methodologies where clinical experts were only presented with KM data from the trial and asked to predict survival at different timepoints.

Abbreviations: DSU: Decision Support Unit; **NICE:** National Institute for Health and Care Excellence; **OS:** overall survival; **PFS:** progression free survival; **TSD:** technical support document.

References: [1]Oakley J. E., Ren S., Forsyth J. E., Gosling J. P., Wilson K., Latimer N., Rutherford M. J., Uttley L., Fotheringham J., NICE DSU Technical Support Document 26: Expert elicitation for long-term survival outcomes. 2025. Available from http://www.nicedsu.org.uk [Last accessed 18 Aug 25]. **Acknowledgements:** The authors thank Freya Turner, Costello Medical, for graphic design assistance.

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