


doi: 10.1093/annonc/mdw296
Published online 8 August 2016

Utilisation of the ESMO-MCBS in practice of HTA

It is highly appreciated that the European Society of Medical Oncology (ESMO) has developed a system to assess new oncologic compounds according to their value to patients [1]. Consequently, offering decision-support to those who either want to use the new cancer therapies in clinical practice but cannot keep up-to-date with all therapy options or, alternatively, to those who have to decide whether or not to fund new oncology medicines or exclude from reimbursement due to their low value. This is particularly important with ever-rising prices for new oncology medicines which have increased up to 10-fold in recent years.

Having established a horizon-scanning system for new oncology medicines (HSO) for Austria [2] since 2009, we have extensive experience with the early assessment of newly approved therapies for patients with cancer (n = 59). Until recently, these assessments have not included recommendations. We are now considering an adapted version of the ESMO-Magnitude of Clinical Benefit Scale (MCBS). With the aim of piloting and validating the ESMO-MCBS—as suggested by Hartmann [3]—we—three researchers, blinded to ESMO scores, blinded to each other’s scoring—rated drugs in three indications (colorectal carcinoma, melanoma and lung cancer) which had been assessed by the Austrian HSO programme as well as having been scored by ESMO (n = 11) (Table 1).

In addition, we discussed the ESMO-MCBS in a meeting of the Piperska-group for ‘rational prescribing’, a group of health authority personnel, advisers and academics from across Europe involved with developing models to optimise the managed entry of new medicines [4, 5] and collected comments based on experiences with applying the proposed scores.

Lastly, we compared our scoring with drug-assessments of several countries [6] (supplementary Table S1, available at Annals of Oncology online) and found a good correlation between oncology medicines scored with 1–2 on the ESMO-MCBS scale with oncology medicines not recommended for funding due to a lack of efficacy or poor cost-effectiveness.

We identified several limitations to the current ESMO-MCBS. Some of them have been pointed out and discussed already by others [3, 7, 8]. As a result, we propose adaptations due to perceived limitations which include:

(i) The use of the lower limit of the confidence interval (CI) rather than the point estimate is not only introducing an optimistic perspective but is also systematically favouring drugs with a large CI and therefore a low certainty in results. We find the systematic bias not acceptable and not in concordance with standards for robust data in medical statistics and clinical epidemiology.

(ii) The focus on primary end points even if they are surrogate end points is assigning progression-free survival (PFS) an equal weight compared with the patient-relevant end points of overall survival (OS) and quality of life (QoL). This is not in accordance with health technology assessment standards of using patient-relevant outcomes as opposed to surrogate outcomes, which are most often not validated for their actual clinical relevance. This is especially important in solid tumours with concerns with translating surrogate markers into overall survival/length of survival.

(iii) The upgrading due to increased QoL, but only rarely the downgrading of cancer drugs due to worsened QoL or (S)AE ≥ 3 (serious adverse events) and/or increased discontinuation of therapy introduces again a bias towards an optimistic perspective concentrating on efficacy and ignoring risks and adverse events.

(iv) Lastly, no rationale or weighted arguments are provided for the threshold values.

As seen in Table 1, our scores deviate from ESMO due to using the point estimate instead of using the lower limit of the CI. In addition, degrading because of (S)AE or if only data on OS are available, rather than only upgrading because of an improvement in QoL, was done. Our re-calculated deviation is on average 1–2 scores lower than the ESMO scoring. We tried to extract as many data as possible in the table to show that even using the rationale of ESMO-MCBS, some of the (optimistic) ESMO-scores are not based on their own rules (alone) and some oncology medicines are upgraded without transparent reason.

We therefore propose an adapted use of the ESMO-MCBS. This includes:

• use of only OS data—if primary or secondary end point data are available;
• use of only OS and QoL data in the non-curative (end of life) setting;
• downgrade −1, if only data on surrogate end points are available;
• use point estimates;
• down-/upgrade due to toxicity(S)AE including reduction in OS or changes in QoL;
• increase transparency with extracting SAE and therapy discontinuation data and reasons for up-/downgrading.

A structured and systematic approach that can discriminate between oncology medicines of higher value than others is most welcomed for assisting in the rational and appropriate use of limited public resources to deliver effective and affordable care. This is becoming more essential with increased prevalence rates worldwide.
<table>
<thead>
<tr>
<th>Active substance (trial name)</th>
<th>Indication</th>
<th>Intention</th>
<th>PE (SE)</th>
<th>Form</th>
<th>MG standard treatment</th>
<th>Efficacy</th>
<th>HR (95% CI)</th>
<th>Score calculation</th>
<th>PM</th>
<th>Safety Toxicity</th>
<th>QoL</th>
<th>Adjustments</th>
<th>HSO</th>
<th>ESMO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table 1. Adapted benefit assessment based on ESMO-MBSC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aflibercept (VELOUR)</td>
<td>mCRC (second line)</td>
<td>Not curative</td>
<td>OS</td>
<td>2a</td>
<td>&gt;1 year</td>
<td>1.44</td>
<td>0.817 (0.71–0.94)</td>
<td>HR &gt;0.75 or OS &lt;1.5 months</td>
<td>1</td>
<td>+21% ≥grade 3 AE; ~15% AEs lead to discontinuation (~1)</td>
<td>–</td>
<td>–</td>
<td>1 (0)</td>
<td>1</td>
</tr>
<tr>
<td>Regorafenib (CORRECT)</td>
<td>mCRC</td>
<td>Not curative</td>
<td>OS</td>
<td>2a</td>
<td>≤1 year</td>
<td>1.4</td>
<td>0.77 (0.64–0.94)</td>
<td>HR &gt;0.75 or OS &lt;1.5 months</td>
<td>1</td>
<td>+21% ≥grade 3 AE; 76% dose reduction or interruption due to AE (~1)</td>
<td>NO impr.</td>
<td>–</td>
<td>1 (0)</td>
<td>1</td>
</tr>
<tr>
<td>Nab-Paclitaxel (MPACT)</td>
<td>m adenocarcinoma of the pancreas (first line)</td>
<td>Not curative</td>
<td>OS</td>
<td>2a</td>
<td>≤1 year</td>
<td>1.8</td>
<td>0.72 (0.62–0.83)</td>
<td>HR &gt;0.70 or OS &lt;1.5 months</td>
<td>1–2</td>
<td>+1–16% ≥grade 3 AE; ~20% AEs lead to discontinuation (~1)</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Afatinib (LUX3)</td>
<td>Locally advanced/mNSCLC (first line)</td>
<td>Not curative</td>
<td>PFS (OS)</td>
<td>2a</td>
<td>&gt;1 year</td>
<td>1.8</td>
<td>OS: 1.12 (0.73–1.73)</td>
<td>HR &gt;0.70–0.75 or OS ≥1.5–2.9 months</td>
<td>2</td>
<td>+1% ≥grade 3 AE; +4% discontinuation</td>
<td>impr QoL: +1</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Afatinib (LUX6)</td>
<td>Locally advanced/ mNSCLC (first line)</td>
<td>Not curative</td>
<td>PFS (OS)</td>
<td>2b</td>
<td>≤6 months</td>
<td>5.4 (–0.1)</td>
<td>PFS: 0.28 (0.20–0.39)</td>
<td>HR ≤0.65 and PFS ≥1.5 months</td>
<td>3</td>
<td>grade5: −0.5; −33.9% discontinuation (~1)</td>
<td>+1/−1b</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Crizotinib (Profile 1007)</td>
<td>NSCLC (ALK positive) (&gt;1 therapy)</td>
<td>Not curative</td>
<td>PFS (OS)</td>
<td>2b</td>
<td>≤6 months</td>
<td>4.7 (not reached)</td>
<td>PFS: 0.49 (0.37–0.64)</td>
<td>HR ≤0.65 and PFS ≥1.5 months</td>
<td>3</td>
<td>+1–14% ≥grade 3 AE; +11% grade 5 (~1)</td>
<td>benefit,IG +1</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Erlotinib (OPTIMAL, CTONG-0802)</td>
<td>NSCLC (first line)</td>
<td>Not curative</td>
<td>PFS (OS)</td>
<td>2b</td>
<td>&gt;1 year</td>
<td>8.5 (not reached)</td>
<td>PFS0.16 (0.10–0.26)</td>
<td>HR &gt;0.70–0.75 or OS ≥1.5–2.9 months</td>
<td>3</td>
<td>−12% SAE (~1)</td>
<td>impr. QoL: +1/−1i</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Erlotinib (EURTAC)</td>
<td>NSCLC (first line)</td>
<td>Not curative</td>
<td>PFS (OS)</td>
<td>2a</td>
<td>&gt;1 year</td>
<td>4.5 (4.1)</td>
<td>OS: 0.80 (0.47–1.37)</td>
<td>HR &gt;0.70–0.75 or OS ≥1.5–2.9 months</td>
<td>3</td>
<td>+4% grade 5 ≤25% ≥grade 3 AE (+1)</td>
<td>–</td>
<td>+1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Gefitinib (IPASS)</td>
<td>Locally advanced/m NSCLC (EGFR) (first line)</td>
<td>Not curative</td>
<td>PFS (OS)</td>
<td>2a</td>
<td>&gt;6 months</td>
<td>3.2 (–0.3)</td>
<td>OS: HR 1.00 (0.76–1.33)</td>
<td>HR &gt;0.70 or gain ≤1.5 months</td>
<td>1</td>
<td>+1,1% grade 5 −32.3% ≥grade 3 AE (+1)</td>
<td>Impr. QoL: +1/−1i</td>
<td>1–2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Iplimumab (NN)</td>
<td>Advanced and mMelanoma (first line)</td>
<td>Not curative</td>
<td>OS</td>
<td>2a</td>
<td>≤1 year</td>
<td>2.1</td>
<td>0.72 (0.59–0.87)</td>
<td>HR &gt;0.65–0.70 or OS &lt;1.5–2.4 months</td>
<td>2</td>
<td>+28% ≥grade 3 AE; Grade (~1)</td>
<td>–</td>
<td>−1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Nivolumab (CheckMate66)</td>
<td>Inoperable/metastatic melanoma (first line)</td>
<td>Not curative</td>
<td>OS (PFS)</td>
<td>2b</td>
<td>≤6 months</td>
<td>30.8% (2.9)</td>
<td>PFS: 0.43 (0.34–0.56)</td>
<td>HR ≤0.65 and PFS ≥1.5 months</td>
<td>3</td>
<td>−5.4% ≥grade 3 AE</td>
<td>2d−1b</td>
<td>−1i</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Trametinib (METRIC)</td>
<td>Advanced/mMelanoma (BRAF V600 mutation)</td>
<td>Not curative</td>
<td>OS (PFS)</td>
<td>2b</td>
<td>≤1 year</td>
<td>3.8% (0.33–0.63)</td>
<td>PFS: 0.45 (0.27–0.63)</td>
<td>HR ≤0.65 and PFS ≥1.5 months</td>
<td>3</td>
<td>+8% ≥grade 3 AE</td>
<td>2d−1</td>
<td>−1b</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Vemurafenib (BRIM-3)</td>
<td>Melanoma (BRAF V600 mutation) (first line)</td>
<td>Not curative</td>
<td>OS</td>
<td>2a</td>
<td>≤1 year</td>
<td>3.3</td>
<td>0.70 (0.57–0.87)</td>
<td>HR &gt;0.65–0.70 OR OS ≥1.5–2.4 months</td>
<td>2–3</td>
<td>≥grade 3 AE +18% ≤cSCC (+5)</td>
<td>–</td>
<td>−1</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

AE, adverse events; benef, benefit; cSCC, cutaneous squamous cell carcinoma; HR, hazard ratio; IG, intervention group; impr, improvement; MG, median gain; OS, overall survival; PE, primary end point; PFS, progression-free-survival; PM, preliminary magnitude of clinical benefit grad; SAE, serious adverse events; SE, secondary end point; QoL, quality of life.

*aMedian OS data from updated analysis; form 2a was used, since OS data available.
*bOnly PFS data, OS not available.
*cOnly PFS data, OS not available.
*dForm 2a was used, since OS data available.
*eForm 2a was used, since OS data available.
*fReduction in OS.
*gForm 2b was used, although OS was the primary end point, but at the time of analysis, it had not been reached.
*hAdjustment (d) Downgrade 1 level if the drug ONLY leads to improved PFS, QoL assessment does not demonstrate improvement.
*iOnly PFS data, OS not available.
C. Wild1, N. Größmann1, P. V. Bonanno2, A. Bucsjics3, J. Furst4, K. Garuolviene5, B. Godman6,7, J. Gulbinovič6, J. Jones9, M. K. Pomorski10 & R. Emprechtlinger1

1Ludwig Boltzmann Institute for Health Technology Assessment (LBI-HTA), Vienna, Austria; 2Independent Pharmaceutical Consultant, Mellieha, Malta; 3Department of Finance, University of Vienna, Austria; 4Medicinal Products Department, Health Insurance Institute of Slovenia, Ljubljana, Slovenia; 5Medicines Reimbursement Department, National Health Insurance Fund, Vilnius, Lithuania; 6Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, UK; 7Department of Laboratory Medicine, Division of Clinical Pharmacology, Karolinska Institutet, Karolinska University Hospital Huddinge, Stockholm, Sweden; 8Faculty of Medicine, Department of Pathology, Forensic Medicine and Pharmacology, Vilnius University and State Medicines Control Agency, Vilnius, Lithuania; 9Healthcare Improvement Scotland, Scottish Medicines Consortium, Glasgow, UK; 10Agency for Health Technology Assessment and Tariff System, Warsaw, Poland

E-mail: claudia.wild@hta.lbg.ac.at

Disclosure

The authors have declared no conflicts of interest.

References


doi: 10.1093/annonc/mdw297
Published online 8 August 2016

Reply to the letter to the editor ‘Utilisation of the ESMO-MCBS in practice of HTA’ by Wild et al.

Dr Wild and colleagues from the Piperska group for ‘rational prescribing’ [1] suggest four limitations in the ESMO-MCBS v1.0. We will address each of them:

1) Regarding the use of lower limit of the 95% CI for HR rather than the point estimate: Use of the point estimate is attractive for its simplicity, but it is statistically problematic. When evaluating for null effect, the upper limit of the 95% confidence interval for HR is used. As an example, for HR 0.75, 95% CI (0.56–1.01), the conclusion would be that at significance level of 0.05, the null hypothesis cannot be rejected. In other words, an HR = 1 cannot be excluded since it is within the 95% CI. This standard is conventional practice for evaluating null effect. No one reasonably insists that, because of the risk of under estimating benefit with the upper limit of the 95% CI, null effect be determined by point estimate ≥1.

Evaluating for magnitude of benefit is statistically analogous. When evaluating magnitude of benefit, the lower limit of the 95% CI for evaluation of benefit indicates that the observed data are congruent with a benefit as great as the target threshold; consequently, insistence on using the point estimate would inappropriately understate the statistically significant potential benefit of the treatment.

The concern that small studies generate wider confidence intervals is real and justified; however, in the ESMO-MCBS v1.0, all high grading scores in a non-curative setting incorporate both HR and absolute gain to mitigate against over valuing small studies with wide HR. Instance on statistical nuance is important vis-a-vis validity and accountability for reasonableness in public health decision-making [2, 3]. These assertions will be further addressed in an upcoming statistical modelling manuscript.

2) Regarding the concern that PFS and OS are equally weighted in v1.0: This is inaccurate. In contrast to OS that can achieve a maximal score of 3/5, preliminary benefit accrued on the basis of PFS is capitated at a maximal score of 3/5, unless there is evidence of substantial survival advantage, improved quality-of-life or diminished toxicity in which case scores can be upgraded [2].

3) Regarding the concern that ESMO-MCBS v1.0 is biased towards an optimistic perspective concentrating on efficacy and ignoring risks and adverse events: The scale incorporates a stratified approach to toxicity penalties; when benefits accrued include improved OS, this version of the MCBS does not incorporate penalty for increased toxicity; when however the score is based on PFS, penalties with negative scoring are introduced both for increased toxicity and also for failure to improve quality of life for failure to delay deterioration in quality of life.

4) Regarding the concern that threshold values in the ESMO-MCBS v1.0 are inadequately justified: As described in our manuscript [2], the threshold values were clinically derived, peer-reviewed for reasonableness and were refined further with statistical modelling and extensive field testing. Furthermore, the thresholds derived in this process were highly consistent with ASCO recommendations for new thresholds for the approval of cancer medications [2, 4].

While acknowledging the potential limitations of the toxicity penalties in v1.0 (which are currently under review), we perceive major substantial shortcomings in at least two of the adaptions proposed by the authors: