EHA evaluation of the ESMO—Magnitude of Clinical Benefit Scale version 1.1 (ESMO-MCBS v1.1) for haematological malignancies

Barbara Kiesewetter,1 Nathan I Cherny,2 Nicolas Boissel,3,4 Francesco Cerisoli,5 Urania Dafni,6,7 Elisabeth G E de Vries,8 Paolo Ghia,9,10 Nicola Gökbüget,11 Verónica González-Calle,12 Brian Huntly,13 Ulrich Jäger,14 Nicola Jane Latino,15 Jean-Yves Douillard,15 Luca Malcovati,15,16,17 Maria-Victoria Mateos,12 Gert J Ossenkoppele,18 Kimmo Porkka,19 Markus Raderer,1 Josep-Maria Ribera,20 Lydia Scarfò,9,10 Ruth Wester,21 Panagiota Zygoura,7 Pieter Sonneveld21,22

ABSTRACT

Objective Value frameworks in oncology have not been validated for the assessment of treatments in haematological malignancies, but to avoid overlaps and duplications it appears reasonable to build up experience on existing value frameworks, such as the European Society for Medical Oncology—Magnitude of Clinical Benefit Scale (ESMO-MCBS).

Methods Here we present the results of the first feasibility testing of the ESMO-MCBS v1.1 for haematological malignancies based on the grading of 80 contemporary studies for acute leukaemia, chronic leukaemia, lymphoma, myeloma and myelodysplastic syndromes. The aims were (1) to evaluate the scorability of data, (2) to evaluate the reasonableness of the generated grades for clinical benefit using the current version and (3) to identify shortcomings in the ESMO-MCBS v1.1 that require amendments to improve the efficacy and validity of the scale in grading new treatments in the management of haematological malignancies.

Results In general, the ESMO-MCBS v1.1 was found to be widely applicable to studies in haematological malignancies, generating scores that were judged as reasonable by European Hematology Association (EHA) experts. A small number of studies could either not be graded or were not appropriately graded. The reasons, related to the differences between haematological and solid tumour malignancies, are identified and described.

Conclusions Based on the findings of this study, ESMO and EHA are committed to develop a version of the ESMO-MCBS that is validated for haematological malignancies. This development process will incorporate all of the usual stringencies for accountability of reasonableness that have characterised the development of the ESMO-MCBS including field testing, statistical modelling, evaluation for reasonableness and openness to appeal and revision. Applying such a scale will support future public policy decision-making regarding the value of new treatments for haematological malignancies and will provide insights that could be helpful in the design of future clinical trials.

What is already known about this subject?
► The European Society for Medical Oncology—Magnitude of Clinical Benefit Scale (ESMO-MCBS) v1.1 is a validated value scale for solid tumour oncology, but it has not yet been evaluated for the use in haematological malignancies.

What does this study add?
► Here, we present the results of the first feasibility testing of the ESMO-MCBS v1.1 for haematological malignancies based on grading of 80 contemporary studies for leukaemia, lymphoma, myeloma and myelodysplastic syndromes.
► The ESMO-MCBS v1.1 was found to be widely applicable to studies in haematological malignancies, generating scores that were judged as reasonable by European Hematology Association (EHA) experts; however, a small number of studies could either not be graded or were not appropriately graded because of shortcomings related to the differences between haematological and solid tumour malignancies.

How might this impact on clinical practice?
► Based on the findings of this study, ESMO and EHA are committed to develop a version of the score that is robustly validated to grade studies in malignant haematology.

INTRODUCTION

In recent years, rapid developments in haematology research resulted in a considerable expansion of treatment options. The development of instruments to measure clinical benefit is essential in the current scenario where increasing numbers of treatments for haematological malignancies (HMs) are becoming available, often targeting a small and defined subpopulation of patients.
For this, several value frameworks have been published by different organisations and institutions taking into account or emphasising different aspects contributing to such an evaluation. These frameworks vary in terms of their definition of value, target audience and methodology, and each of them has specific limitations, which should be taken into consideration when interpreting their outputs. Until now, value frameworks developed in oncology have not been validated in the setting of HMs.

The European Society for Medical Oncology (ESMO) has developed such a value framework called the ESMO—Magnitude of Clinical Benefit Scale (ESMO-MCBS). Initially published in 2015, the scale is a validated and reproducible tool in solid tumour oncology with a particular focus on the clinical benefit. The ESMO-MCBS was developed to generate clear, valid and unbiased grading of the magnitude of clinical benefit demonstrated in therapeutic studies that could be used for a number of purposes including public health policy and health technology assessment (HTA), clinical decision-making, medical publication and journalism. The ESMO-MCBS grading highlights those treatments which substantially improve the duration of survival and/or the quality of life (QOL) of patients with cancer and aims to distinguish them from trials demonstrating more limited and sometimes even marginal benefits. The ESMO-MCBS was revised (version 1.1) in 2017, based on feedback and queries from clinicians, patients, researchers and representatives of the pharmaceutical industry, and a dynamic process of internal peer review. Version 1.1 incorporates 10 revisions and most importantly allows also for scoring of single-arm studies. The ESMO-MCBS assigns categorical benefit scores to European Medicines Agency (EMA) approved drugs, based on results from ‘positive’ randomised clinical trials: (1) superiority trials that have demonstrated a statistically significant result for the primary endpoint of the study, or secondary in case of overall survival (OS) and (2) non-inferiority trials, reaching a conclusion of non-inferiority. Primary or secondary endpoints included in the scoring system are OS, progression-free survival (PFS), QOL, treatment toxicity or response rates. In developing the ESMO-MCBS scale, ESMO aspired to meet standards for ‘accountability for reasonableness’, incorporating extensive field testing, statistical modelling and peer review of the reasonableness of the generated results into the development process. The ESMO-MCBS is currently incorporated in ESMO’s clinical practice guidelines and is being used as part of HTA processes.

The European Hematology Association (EHA) and ESMO have developed a joint initiative to develop a version of the ESMO-MCBS that is validated for HMs. As a first step in this process, we have field tested the current version of the ESMO-MCBS (version 1.1) across a wide spectrum of HMs. The aims of this evaluation were (1) to evaluate the scorability of data derived from contemporary clinical trials in HMs, (2) to evaluate the reasonableness of the generated grades for clinical benefit using the current version and (3) to identify shortcomings in the ESMO-MCBS v1.1 that require amendments to improve the efficacy and validity of the scale in grading new treatments in the management of HMs.

METHODS
Study selection

The corresponding disease-oriented EHA scientific working groups identified experts who selected representative treatments currently used in clinical practice with a focus on recently approved drugs and novel strategies, to be evaluated for each of the common haematological malignancies: acute lymphoblastic leukaemia (ALL), acute myeloid leukaemia (AML), chronic lymphocytic leukaemia (CLL), chronic myeloid leukaemia (CML), Hodgkin and non-Hodgkin lymphomas, multiple myeloma (MM) and myelodysplastic syndromes (MDS). The treatments selected underwent a literature search to identify corresponding clinical trials and data.

ESMO-MCBS grading

Identified studies were graded by members of the EHA scientific working groups according to the ESMO-MCBS v1.1 forms in accordance with the instructions provided by ESMO. Magnitude of clinical benefit scores range from A to C for treatment strategies with curative intent and 5-1 for treatments with non-curative intent, with scores of A–B and 5-4 relating to a substantial level of clinical benefit. Initial grading by the expert groups were reviewed by the ESMO-MCBS working group for applicability and correctness.

Evaluations

For each disease entity, we evaluated the scorability of the evaluated studies and the reasonableness of the derived scores. Based on these findings, we identified shortcomings in the current version of the ESMO-MCBS that either precluded scoring or which generated grading which was considered not to be a reasonable estimation of benefit when such studies were identified.

RESULTS

The extensive research concluded in 80 studies, 5 of which had either more than two arms or different publications for the same trial presenting results after longer follow-up times (87 studies and/or comparisons in total). In detail, we have scored 7 studies for AML, 5 studies for ALL, 8 studies for CLL, 4 studies for CML, 23 studies for non-Hodgkin and Hodgkin lymphoma, 23 studies for MM and 10 studies for MDS. The ESMO-MCBS v1.1 tool was applied in all the 87 distinct studies and/or subgroups.

Acute myeloid leukaemia

Studies evaluated. Seven studies were evaluated three in a curative setting and four in a non-curative setting (table 1).
Scorability: All studies were published with endpoints and data applicable to the ESMO-MCBS v1.1.

Reasonableness: The separation of studies with curative/non-curative intent corresponds closely to the distinction between intensive versus non-intensive chemotherapy regimens which are the terms usually applied in the treatment of AML. Grading effectively distinguished between high benefit treatment strategies in a curative setting and stratified between higher and lower benefit treatments in a non-curative setting.

Shortcomings: None identified.

Acute lymphoblastic leukaemia

Studies evaluated: Five studies were evaluated, and these included studies relating to three agents recently approved by EMA for relapsed and refractory ALL (table 2). 

Scorability: Four of the five studies were published with endpoints and data applicable to the ESMO-MCBS v1.1. The only not scoreable study was the single-arm study of ponatinib as add-on to standard of care upfront treatment with curative intent. 

Reasonableness: Both the first-in-class bispecific antibody blinatumomab (TOWER trial) and the antibody-drug conjugate inotuzumab ozogamicin (INO-VATE trial) reached high scores based on positive OS data and favourable QOL data for blinatumomab (ESMO-MCBS v1.1 scores 5 and 4, respectively). The chimeric antigen receptor (CAR) T-cell treatment in children/young adults with relapsed or refractory B-cell ALL was graded with maximal credit of 3 for a single-arm study in a non-curative setting. The ponatinib treatment (single-arm PACE trial) was assigned grade 2 based on the major molecular response (MMR) in the non-curative setting.

Reasonableness: Grading effectively distinguished between high benefit treatment strategies in a curative setting and stratified between higher and lower benefit treatments in a non-curative setting.

Shortcomings: One shortcoming was identified:

1. The ESMO-MCBS v1.1 does not have a form to grade single-arm treatments with curative intent. This shortcoming precluded scoring in one study and may also have been relevant to the grading of CAR T-cell salvage therapy which could also be considered as curative.

Chronic lymphocytic leukaemia

Studies evaluated: Eight studies were evaluated (table 3).

Scorability: CLL is generally a relatively indolent disease with a very long survival—often decades long—and many patients do not need intervention for many years and when treatment is initiated it commonly generates very long periods of remission. For these reasons, PFS is generally the most relevant and measurable primary endpoint. Since CLL is generally not considered to be a curable disease, all scoring was performed using scales for non-curative disease. One study could not be scored because the primary objective of non-inferiority with regard to PFS was not met. Moreover, the published results limited to a
| Medication                        | Trial name         | Setting                      | Primary outcome | PFS/EFS control | PFS/EFS gain | OS control | OS gain | OS HR | RR (DOR) | QOL | Toxicity | ESMO-MCBS score | ESMO-MCBS form | Reference(s) |
|----------------------------------|--------------------|------------------------------|-----------------|------------------|--------------|-------------|----------|-------|----------|-----|----------|----------------|----------------|--------------|
| Blinatumomab versus SOC          | TOWER              | Relapsed/refractory          | OS              | 12% EFS          | 19%          | 4 months    | 3.7 months| 0.71  | (0.55–0.93) |     | 44% vs 25% CRR, gain 19% | Improved (+1 point) |              | 5, 2a       |
| Inotuzumab ozogamicin versus SOC | INO-VATE           | Relapsed/refractory          | OS/CRR          | 1.8 months       | 3.2 months   | 6.7 months  | 13% gain | 0.77  | (97.5% CI: 0.58 to 1.03) |     | 81% vs 29% CRR, gain 52% | Improved Veno-occlusive disease 11% in experimental arm | 4*, 2a      |
|                                  |                    |                              | EFS             | 81% 2 years EFS |              | 80% 2 years |          |       |          |     | Not scoreable |                |                |             |
| Hyper-CVAD + ponatinib           |                    |                              | EFS             | 81% 2 years EFS |              | 80% 2 years |          |       |          |     | Not scoreable |                |                |             |
| CAR T-cell                     |                    |                              | EFS             |                  |              | 81% ORR    |          |       |          |     | >30% grade 3/4 cytokine release syndrome |                | 3, 3        |
| tisagenlecleucel                |                    |                              | EFS             |                  |              | 81% ORR    |          |       |          |     | >30% grade 3/4 cytokine release syndrome |                | 3, 3        |
| Ponatinib                        |                    |                              | EFS             | Major haematological response within the first 6 months | 40% at 12 months | Major haematological response: 41% (3 months) |          |       |          |     | Not scoreable |                |                |             |

*Based on >10% increase in 2 years of OS improvement.

CAR T-cell, chimeric antigen receptor T-cell therapy; ORR, complete remission rate; DOR, duration of response; EFS, event-free survival; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; Hyper-CVAD, hyperfractionated cyclophosphamide, vincristine, doxorubicin and dacarbazine; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RR, response rate; SOC, standard of care; TKI, tyrosine kinase inhibitor.
### Table 3  Feasibility testing of the ESMO-MCBS v1.1 for chronic lymphocytic leukaemia (n=8)

| Medication       | Trial name      | Setting                                | Primary outcome | PFS control | PFS gain | PFS HR       | OS control | OS gain | OS HR       | RR       | QOL          | Toxicity                                                                 | ESMO-MCBS score | ESMO-MCBS form | Reference(s) |
|------------------|-----------------|----------------------------------------|-----------------|-------------|----------|---------------|-------------|---------|---------------|----------|----------------|--------------------------------------------------------------------------|-----------------|----------------|-------------|
| FC±R             | CLL8            | Upfront, chemofit                      | PFS             | 32.9 months| 23.9 months| 0.59 (0.50–0.69)| 86 months   | >10% gain at 8 years | 0.68 (0.54–0.89) | No difference | Increased |                                                                                     | 4               | 2a             | 24-26        |
| FC-R versus R- bendamustine | CLL10          | Upfront, focus elderly subgroup >65 years | Non-inferiority in PFS | 55.2 months | −13.5 months | Non-inferiority not met neither overall, nor in the >65 years post hoc subgroup | Not significant |                                                                                     | 2c               |                                                                                     |
| Ibrutinib versus chlorambucil | RESONATE-2     | Upfront elderly                        | PFS             | 18.9 months| 8 months  | 0.16 (0.09–0.26)| 85% at 12 months| 13%       | 0.16 (0.05–0.56) | Immature | Improved (abstract only) |                                                                                     | 3               | 2b             | 26-29        |
| Obinutuzumab± chlorambucil | CLL11          | Upfront elderly not eligible for fludarabine | PFS             | 11.1 months| 15.6 months| 0.18 (0.13–0.24) | NR         | NA       | 0.41 (0.23–0.74) | Immature | Increased but not meeting criteria for downgrading |                                                                                     | 3               | 2b             | 30          |
| Ibrutinib versus ofatumumab | RESONATE | Relapsed/refractory (cross-over allowed) | PFS             | 8.1 months | 4+ months (≥10% gain at 12 months with plateau) | 0.11 (0.03–0.15) | 81% at 12 months | 9% at 12 months | 0.43 (0.24–0.79) | Immature | Pending >10% SAE increase (≥1 point) |                                                                                     | 3               | 2b             | 31-32        |
| R-Venetoclax versus R- bendamustine | MURANO        | Relapsed/refractory                    | PFS             | 17 months  | 6+ months (≥10% gain at 12 months with plateau) | 0.17 (0.11–0.26) | 87% at 12 months | 5.30%      | 0.48 (0.25–0.70) | Immature |                                                                                     | 4               | 2b             | 33          |
| Ibrutinib        | RESONATE-17     | Relapsed/refractory                    | ORR             | 63% at 24 months | 75% at 24 months | 64% |                                                                                     | 3               | 3             | 34          |
| Venetoclax       | M13-982         | Relapsed/refractory                    | ORR             | 72% at 12 months | 87% at 12 months | 79% |                                                                                     | 3               | 3             | 35          |

del17p, 1p deletion; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; FC, fludarabine, cyclophosphamide, NA, not applicable; NR, not reached; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; QOL, quality of life; R, rituximab; RR, response rate; SAE, serious adverse event.
subcohort of patients older than 65 years, which are relevant for clinical practice (particularly in view of presented toxicity data) did not show non-inferiority and they were derived from a post hoc exploratory analysis.

**Reasonableness** Overall scoring was considered reasonable with the highest grades being achieved by studies demonstrating either mature OS data or PFS gains with long-term plateauing of PFS or compelling PFS gains. Grading of the phase III study of ibrutinib versus ofatumumab (RESONATE trial) was considered to be low; it was credited for PFS advantage including gain in the tail of the curve but was penalised for toxicity associated with the more prolonged drug exposure in continuous treatment (ESMO-MCBS v1.1 score 3). However, the 9% improvement in OS at 12 months was not credited as these results are deemed immature by the ESMO-MCBS criteria. The benefit of novel agents in populations with high unmet need, like relapsed and refractory patients with CLL carrying deletion in chromosome 17 p, was graded reasonably using form 3 for single-arm studies in a non-curative setting.

**Shortcomings:** One shortcoming was identified:
1. The EHA scientific working group members felt that compelling immature survival benefit ought to be credited even when the median survival of the control arm has not been reached.

**Chronic myeloid leukaemia**

**Studies evaluated:** Four landmark trials addressing the use of tyrosine kinase inhibitors imatinib, nilotinib, dasatinib and bosutinib upfront for chronic phase CML were graded. Only one of these had mature OS data (table 4).

**Scorability:** CML is generally considered an incurable disease, but in a small proportion of cases with deep molecular responses the disease may be eradicated. Thus, when mature survival data were available, CML was scored for both curative and non-curative intent. Contemporary studies in CML treatments are conventionally evaluated using molecular response evaluations. This differs from the concepts of ‘pathological complete response’ or ‘response rate’ which are terms used in the ESMO-MCBS v1.1. Scoring of these studies was only possible by interpreting deep molecular responses (MMR 4–5) as pathological complete responses (form 1) or major responses (form 2c). In one study, PFS/event-free survival (EFS) gains could not be credited because the PFS of the control arm was very long and had not reached median PFS after 11 years of follow-up.

**Reasonableness:** In the IRIS study of imatinib versus former standard interferon plus cytarbine, initial scoring at 18 months was credited on the basis of complete cytogenetic response for curative intent with a grade of C and improvement in molecular response rate with grade 2. At 10-year follow-up, the imatinib scores B for curative intent based on survival improvement. While the grades for curative intent were considered reasonable, the EHA working group considered the ESMO-MCBS grade of 2 for non-curable intent to be too low for the benefits observed.

The remaining studies of nilotinib, dasatinib and bosutinib show minor improvements in complete molecular response rates when compared with imatinib (grade 2) in a non-curative setting. None of these agents had mature data beyond 5 years and consequently they were not graded for curative intent.

**Shortcomings:** These relatively low scores for imatinib in the non-curative grading appear to indicate two shortcomings in the ESMO-MCBS v1.1:
1. When PFS (or EFS) is very long, there is no mechanism to credit strong interim gains when the median PFS of the control arm has not yet been reached.
2. The surrogacy of complete cytogenic response and level 4–5 MMR, defined as 4 to 5-log reduction in BCR–ABL1 transcript levels from a standardised baseline, are much stronger surrogates for survival than pathological complete response and response rate in solid tumours. Consequently, form 2c needs to be amended to incorporate evaluation of deep molecular responses.

**Indolent non-Hodgkin’s, relapsed/refractory setting of non-diffuse large B-cell lymphoma (non-DLBCL) and Hodgkin’s lymphoma**

**Studies evaluated:** Twelve studies of recently approved drugs for indolent non-Hodgkin’s, relapsed/refractory setting of non-DLBCL and Hodgkin’s lymphoma were evaluated (table 5).

**Scorability:** In one of the studies, PFS/EFS gains could not be graded because the PFS of the control arm was very long, the median PFS was not reached and only interim gains were reported. The BRIGHT study could not be scored because form 2c makes no provision for scoring of non-inferiority studies based on response rates. The remaining 10 studies were published with endpoints and data applicable to the ESMO-MCBS v1.1 and were all evaluable.

**Reasonableness:** The grading was applicable and was judged by the EHA working group to be reasonable in the evaluated trials, endorsing relatively high benefit grades, that is, ESMO-MCBS v1.1. scores of 4–5 for 7 of the 10 evaluable studies.

**Shortcomings:** Two shortcomings were observed:
1. The ESMO-MCBS v1.1 has no mechanism for scoring non-inferiority studies based on response rate.
2. When PFS (or EFS) is very long, there is no mechanism to credit strong interim gains when the median PFS of the control arm has not yet been reached.

**Diffuse large B-cell lymphoma**

**Studies evaluated:** Eleven studies were evaluated, two in the first-line setting with curative intent, two intensified therapies for first-line and salvage setting, respectively, with both curative and non-curative intent, two single-arm studies of CAR T-cell salvage therapy and five in a non-curative setting for relapsed and refractory disease (table 6).
### Table 4 Feasibility testing of the ESMO-MCBS v1.1 for chronic myeloid leukaemia (n=4)

| Medication                          | Trial name      | Setting                                                                 | Primary outcome                                                                 | EFS/PFS control | EFS/PFS gain | OS control | OS gain | OS HR | Improved | Less toxicity | Toxicity | ESMO-MCBS score | ESMO-MCBS form | Reference(s) |
|-------------------------------------|-----------------|-------------------------------------------------------------------------|---------------------------------------------------------------------------------|------------------|--------------|------------|---------|------|----------|---------------|----------|----------------|----------------|--------------|
| Imatinib versus interferon/cytarabine | IRIS            | Newly diagnosed chronic phase (cross-over allowed)                      | Initial: PFS/EFS long term: OS                                                 | 18 months PFS    | 73.5%        | 18.6%      |         |      |          |               |          |                |                | C/2          |
|                                     |                 |                                                                         | 10 years EFS                                                                   | 56.6%            | 23%          | 78.8%      | 4.5%    | 0.74  | 0.66-0.99|               |          |                |                | B/2 2c       |
| Nilotinib 600 or 800 mg versus imatinib | ENESTnd         | Newly diagnosed chronic phase                                           | Initial primary: MMR at 12 months, secondary: complete cytRR                    | 12 months 600mg  | 80% vs 65%,  gain 15% | 44% vs 22%, gain 22% |         |      |      |          |               |          |                |                | 2 2c         |
|                                     |                 |                                                                         | 12 months 800mg                                                                | 78% vs 65%, gain 13% | 43% vs 22%， gain 21% |         |      |      |          |               |          |                |                | 2 2c         |
|                                     |                 |                                                                         | 5 years 600 mg                                                                 | 92.6%            | 2.4%         | NS         | 91.7%   | 2.0%  | NS       |               |          |                |                | 2 2c         |
|                                     |                 |                                                                         | 5 years 800 mg                                                                 | 4.3%             | 0.37 (0.15-0.88) | 4.5%       | 0.44    | 0.74  | 0.56-0.99|               |          |                |                | 2 2c         |
| Dasatinib versus imatinib           | DASISION        | Newly diagnosed chronic phase                                           | Complete cytRR                                                                 | 12 months        | 77% vs 66%, gain 11% | 46% vs 28%, gain 22% |         |      |      |          |               |          |                |                | 1 2c         |
|                                     |                 |                                                                         | 5 years                                                                         | 90%              | 1%           | NS         |         |      |          |               |          |                |                | 2c           |
| Bosutinib versus imatinib           | BEFORE          | Newly diagnosed chronic phase                                           | MMR at 12 months                                                               | 77% vs 66%, gain 11% | 47% vs 37%, gain 10% |         |      |      |          |               |          |                |                | 1 2c         |

Cardiovasc., cardiovascular; CytRR, cytogenetic response rate; EFS, event-free survival; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; MMR, major molecular response; MR, molecular response; NS, not significant; OS, overall survival; PFS, progression-free survival; QOL, quality of life.
### Table 5  Feasibility testing of the ESMO-MCBS v1.1 for indolent non-Hodgkin and relapsed/refractory setting of non-DLBCL and Hodgkin’s lymphoma (n=12)

| Medication | Trial name | Setting | Primary outcome | PFS control | PFS gain | PFS HR | OS control | OS gain | OS HR | RR (DOR) | QOL | Toxicity | ESMO-MCBS score | ESMO-MCBS form | Reference(s) |
|------------|------------|---------|----------------|--------------|----------|--------|------------|---------|-------|----------|-----|----------|----------------|----------------|--------------|
| Obinutuzumab-Chemo versus R-Chemo | GALLIUM | Follicular lymphoma, first line | PFS | 73% (3 years) | 7% | 0.66 (0.51–0.85) | Not scoreable | 2b | 46 |
| VR-CAP versus R-CHOP | LYM-3002 | Mantle cell lymphoma first line, not eligible for transplant | PFS | 14.4 months | 10.3 months | 0.63 (0.5-0.79) | 95.7 months | 0.66 (0.51–0.85) | Increased in experimental arm | A/4 | 1/2a | 47-48 |
| R-Bendamustine versus R-CHOP/R-CVP | BRIGHT study | Indolent and mantle cell lymphoma, first line | Non-inferiority in PFS (margin: 1.32) | 31.2 months | 38.3 months | 0.58 (0.44–0.74) | Non-inferiority met | Improved | Not scoreable | 2c | 49-50 |
| R-Bendamustine versus R-CHOP | STIL Trial NHL 1-2003 | Indolent and mantle cell lymphoma, first line | Non-inferiority in PFS (margin: 0.88) | 14.9 months | NA | 0.55 (0.40–0.74) | Non-inferiority met | Delayed deterioration in QOL | 5 | 2a | 50-54 |
| Bendamustine ± Obinutuzumab | GADOLIN | Rituximab-refractory indolent non-Hodgkin's lymphoma | PFS | 14.9 months | NA | 0.67 (0.47–0.96) | >10% at 5 years | Improved (+1 point) | 4 | 2b | 55 |
| Ibrutinib versus Temsirolimus | RELAPSED study | Relapsed/refractory mantle cell lymphoma | PFS | 6.2 months | 8.4 months | 0.43 (0.32–0.58) | Improved (+1 point) | 4 | 2b | 56 |
| Lenalidomide versus Investigator's choice | MCL-002 SPRINT | Relapsed/refractory mantle cell lymphoma | PFS | 5.2 months | 3.5 months | 0.61 (0.44–0.84) | Improved (+1 point) | 4 | 2b | 57 |
| Ibrutinib | PCY-1104-CA | Relapsed/refractory mantle cell lymphoma | ORR | 13.9 months | 68% | 0.67 (0.47–0.96) | >10% at 5 years | Improved (+1 point) | 3 | 3 | 58 |
| Ibrutinib | Relapsed/refractory marginal zone lymphoma | ORR | 14.2 months | 48% | Relevant toxicity but not meeting criteria for downgrading | 3 | 3 | 59 |
| Idealisib | DELTA (101-009) | Relapsed/refractory indolent lymphoma | ORR | 11 months | 57% (12.5 months) | 3 | 3 | 60 |
| Pembrolizumab | KEYNOTE-087 | Relapsed/refractory Hodgkin lymphoma | ORR | 69% | Improved (+1 point) | 4 | 3 | 61 |
| Nivolumab | Check Mate 205 | Relapsed/refractory Hodgkin lymphoma | ORR | 14.7 months | 69% | Improved (+1 point) | 4 | 3 | 62 |

chemo, chemotherapy; CRR, complete response rate; DLBCL, non-diffuse large B-cell lymphoma; DOR, duration of response; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; NA, not applicable; NR, not reached; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; QOL, quality of life; R, rituximab; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone; R-CVP, rituximab, cyclophosphamide, vincristine and prednisone; RR, response rate; VR-CAP, bortezomib, rituximab, cyclophosphamide, doxorubicin and prednisone.
Table 6 Feasibility testing of the ESMO-MCBS v1.1 for DLBCL (n=11)

| Medication | Trial name | Setting | Primary outcome | PFS/EFS/DFS control | PFS/EFS/DFS gain | PFS/EFS/DFS HR | OS control | OS gain | OS HR | RR (DOR) | QOL | Toxicity | ESMO-MCBS score | ESMO-MCBS form | Reference(s) |
|------------|------------|---------|-----------------|----------------------|------------------|----------------|--------------|---------|--------|--------|---------|-----|----------|----------------|----------------|--------------|
| CHOP±R     | MinT study | First-line DLBCL, stage II–IV or I with bulky disease, IPI 0–1 | EFS | 55.6% (6 years) | 18.5% | p<0.0001 | 80% (6 years) | 10.1% | p=0.0004 | A | 1 | 64 |
| CHOP±R     | LNH-98.5   | First-line DLBCL, stage II–IV, age 60–80 | PFS | 20% at 10 years | 16.5% | p<0.0001 | 27.8% (10 years) | 15.9% | p<0.0001 | A | 1 | 65,66 |
| R-CHOP     | REMARC     | First-line DLBCL, stage II–IV, age 60–80 | PFS | 58.9% at 4+ months | 0.71 (0.54–0.93) | NS | No difference | No difference | Improved | B | not scoreable | 1/2c | 67 |
| R-GDP±ASCT +ASCT | NCIC-CTG LY12 | Relapsed/refractory aggressive lymphoma | PFS | 58.9% (6 years) | 18.5% | p<0.0001 | 80% (6 years) | 10.1% | p=0.0004 | A | 1 | 66 |

**Pixantrone versus investigators’ choice** | Relapsed/refractory aggressive lymphoma | CRR | 2.6 months | 2.7 months | >10% gain at 12 months, no plateau | 0.60 (0.42–0.96) | 20% vs 6%, gain 14% | | 3 | 2b | 69 |

**CAR T-cell** | Axicabtagene ciloleucel | ZUMA-1 | Relapsed/refractory aggressive non-Hodgkin’s lymphoma | ORR | 82% | Toxicity but not meeting criteria for downgrading | | 3 | 3 | 70 |

**CAR T-cell** | Tisagenlecleucel | JULIET | Relapsed/refractory DLBCL | ORR | 52% (not reached, >10 months) | Toxicity not meeting criteria for downgrading | | 3 | 3 | 71 |

**Lenalidomide versus investigators’ choice** | DLC-001 | Relapsed/refractory DLBCL | ORR | 2 months | 1.4 months | 0.64 (0.41–0.99) | 28% vs 12%, gain 16% | More PFS improvement in ABC subtype | 2 | 2b | 72 |

**Panobinostat with or without R** | Relapsed/refractory DLBCL | ORR | 28% (15 months) | 3 | 3 | 73 |

**Brentuximab vedotin** | Relapsed/refractory DLBCL | ORR | 44% | 3 | 2 | 74 |

**Ibrutinib** | Relapsed/refractory DLBCL, subgroup ABC subtype | ORR | 37% (4.8 months) | 1 | 3 | 75 |

ASCT, autologous stem cell transplantation; CART-cell, chimeric antigen receptor T-cell therapy; CHOP, cyclophosphamide, doxorubicin, vincristine and prednisone; ORR, complete response rate; DFS, disease-free survival; DLBCL, diffuse large B-cell lymphoma; DOR, duration of response; EFS, event-free survival; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; IPI, International Prognostic Index; NS, not significant; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RR, response rate; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone; R-DHAP, rituximab, dexamethasone, cytarabine and cisplatin; R-GDP, rituximab, gemcitabine, dexamethasone and cisplatin; RR, response rate.
Scorability: All studies incorporated required data for evaluation using the ESMO-MCBS v1.1. Single-arm studies of CAR T-cell therapy for refractory or resistant disease could not be evaluated for curative intent. The NCIC-CTG LY12 trial could not be graded in the non-curative setting because non inferiority was evaluated on the basis of overall response rate.

Reasonableness: The grading was applicable and was judged by the EHA working group to be reasonable in the evaluated trials, endorsing high benefit grades for first-line therapies with curative intent. Lower benefit scores for trials in the relapsed and refractory therapies were considered reasonable.

Shortcomings: One shortcoming was identified:
1. The ESMO-MCBS v1.1 does not have a form to grade single-arm treatments with curative intent and this shortcoming does not allow for the representation of the full potential benefit of CAR T-cell salvage therapy.

Multiple myeloma

Studies evaluated: Table 7 describes results from eight studies in the first-line setting. Of these, three were conducted for autologous stem cell transplantation (ASCT) eligible patients and five are for ASCT ineligible patients. Table 8 describes the results of a further 15 studies with relapsed or refractory myeloma.

Scorability: Most studies incorporated required data for evaluation using the ESMO-MCBS v1.1. The PETHEMA/GEM study comparing VTD (bortezomib, thalidomide and dexamethasone) to TD (thalidomide and dexamethasone) or VBMCP/VBAD/B (vincristine, BCNU, melphalan, cyclophosphamide, prednisone/vincristine, BCNU, doxorubicin, dexamethasone/bortezomib) as induction therapies did not report HRs for the PFS, resulting in precluded scoring with non-curable intent using form 2b. The GIMEMA 2005 study could not be scored for non-curable intent because the median PFS of the control arm has not yet been reached. The MM5 non-inferiority study could not be scored for non-curable intent because non-inferiority was based on response rate.

Reasonableness: First-line treatments for patients who are ASCT eligible are graded both for curative and non-curable intent. The relatively low grades of C for curative intent achieved in two of the ASCT eligible studies reflect the prevailing consensus that MM is rarely cured. In most studies evaluated, the scale was feasible and the results were consistent with clinical practice.

Shortcomings: Three previously described shortcomings influenced scoring for a small number of these studies.
1. The ESMO-MCBS v1.1 has no mechanism for scoring non-inferiority studies in a non-curative setting based on response rate.
2. When PFS (or EFS) is very long, the ESMO-MCBS v1.1 has no mechanism to credit strong interim gains when the median PFS of the control arm has not yet been reached.
3. The EHA working group members felt that the capitation of PFS at a maximal preliminary grade of 3, with provision for an upgrade based on tail of the curve only if there is a plateau in the study medication PFS with gain of >10% at 12 months, may have undervalued some MM treatments. The plateau requirement for this adjustment precludes credit for substantial prolonged gains in PFS in this disease entity.

Myelodysplastic syndrome

Studies evaluated: Ten studies were evaluated in this setting. Of these, two studies were evaluated based on OS or PFS and the remaining eight studies were evaluated based on response rate (table 9).

Scorability: All studies incorporated required data for evaluation using the ESMO-MCBS v1.1. Clinical benefit measure was, however, partly confounded by the heterogeneity of the available definitions of haematological response and their clinical meaningfulness.

Reasonableness: In the two studies evaluating hypomethylating agents in intermediate-risk/high-risk patients, the ESMO-MCBS v1.1 graded them with substantial benefit based on either PFS gain or OS gain with improved QOL. In lower risk patients, the remaining eight studies included randomised trials investigating erythropoietin-stimulating agents, lenalidomide in MDS with del(5q) or non-del(5q) and immunosuppressive therapy with antithymocyte globulin plus cyclosporine, compared with best supportive care. All studies were evaluated based on response rates, but they used a range of different and inconsistent criteria, some using International Working Group, or modifications thereof, and other study-specific criteria such as transfusion requirements. All these studies resulted in a final ESMO-MCBS v1.1 score of 2. In one of these studies QOL was evaluated and demonstrated to have improved but this was not reflected in grading since there is no QOL bonus for studies in which response rate is the primary outcome.

Shortcomings: The EHA working group identified one shortcoming derived from these evaluations:
1. In studies evaluating response rate as a primary endpoint, there is no provision of QOL bonus if improved QOL is demonstrated as a secondary outcome.

DISCUSSION

The EHA with currently more than 5000 members is the largest European-based haematology association. In addition to its educational mission, it has a public policy and advocacy role that engages stakeholders, including patient representatives, to improve patient care and to raise awareness for haematology as a distinct medical discipline with specific needs. Reflecting these goals, EHA has observed the development of the ESMO-MCBS and its broad utility in solid tumour oncology with great interest, and in the absence of a value tool validated for malignant haematology, we sought to investigate the
### Table 7 Feasibility testing of the ESMO-MCBS v1.1 for first-line multiple myeloma (n=8)

| Medication          | Trial name       | Setting       | Primary outcome | PFS/DFS control | PFS/DFS gain | PFS/DFS HR | OS control | OS gain | OS HR | RR | QOL | Toxicity | ESMO-MCBS score | ESMO-MCBS form | Reference(s) |
|---------------------|------------------|---------------|-----------------|-----------------|--------------|------------|------------|---------|-------|-----|-----|-------|-----------------|----------------|--------------|
| VTD versus TD       | GEM2005-lesss65  | PETHEMA/GEM   | CR post ASCT (PFS) | 28.2 months | 28.0 months | p=0.01 | 65% at 4 years | 9% | NS | CRR 46% vs 24%, gain 22% | More neuropathy but not meeting criteria for downgrading | C/not scoreable | 1/2b | 76 |
| VBMCP/VBAD/B        |                  |               |                 | 35.3 months | 20.9 months | p=0.01 | 70% at 4 years | 4% | NS | CRR 46% vs 38%, gain 8% |                      | NEB/not scoreable |                  |
| VTD versus TD       | GIMEMA 2005      | ASCT eligible | CR post induction (PFS) | 56% at 3 years | 12% | 0.63 (0.45–0.88) | 84% at 3 years | 2% | NS | (near) CRR 31% vs 11%, gain 20% | More neuropathy but not meeting criteria for downgrading | C/not scoreable | 1/2b | 77 |
| VCD versus PAD      | MM5              | ASCT eligible | Non-inferiority of ≥VGPR rate (margin: ≥-10%) |                 |              |              |              |              |      |      |     | VQIPR difference: 2.8% (-6.8% to 12.3%) non-inferiority met | SAE’s higher in the control arm | Not scoreable | 1/2c | 78 |
| VMP versus MP       | VISTA            | ASCT ineligible | TTP | 16.6 months | 7.4 months | 0.48 (p<0.001) | 43.1 months | 13 months | 0.70 (0.67–0.85) |                      |                  | 4 | 2a | 79 80 |
| VMPT versus VMP     | GIMEMA VMPT      | ASCT ineligible | PFS | 27 months | >13 months | 0.67 (0.50–0.80) | 87% at 3 years | 2% | NS | – | Vascular and cardiac events increased in experimental arm (1 point) |                  | 2 | 2b | 81 |
| Lenalidomide-d      | FIRST            | ASCT ineligible | PFS | Len-d x18 | 20.7 months | 4.8 months | 0.70 (0.60–0.82) | 56% at 4 years | 3% gain at 4 years | NS |                  |                  | 3 | 2b |
| continuous versus   |                  |               |                 | MPT | 21.2 months | 4.3 months | 0.72 (0.61–0.80) | 47 months | 51% at 4 years | 8% gain at 4 years | 0.78 (0.66–0.96) |                  |                  | 4 | 2a |
| x18 or MPT x12      |                  |               |                 |                 |              |              |              |              |      |      |     |                  |                  |              |
| VMP ±daratumumab    | ALOYONE          | ASCT ineligible | PFS | 18 months | 50% at 18 months | 9+ months | 21% at 18 months | 0.50 (0.38–0.69) |              |              | 3 | More infections but not meeting criteria for penalty |                  | 3 | 2b |
| ±bortezomib         |                  |               |                 |                 |              |              |              |              |      |      |     |                  |                  |              |
| Lenalidomide-d      | SWOG 50777       | ASCT ineligible | PFS | 30 months |              |              |              | 64 months | 11 months | 0.71 (0.52–0.96) | Slightly increased |                  | 4 | 2a |
| continuous versus   |                  |               |                 |                 |              |              |              |              |      |      |     |                  |                  |              |
| MPT                 |                  |               |                 |                 |              |              |              |              |      |      |     |                  |                  |              |

ASCT, autologous stem cell transplantation; CR, complete remission; CRR, complete remission rate; d, dexamethasone; DFS, disease-free survival; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; Len-d, lenalidomide-d; MP, melphalan and prednisone; MPT, melphalan, prednisone and thalidomide; NEB, no evaluable benefit; NS, not significant; OS, overall survival; PAD, prednisone; PFS, progression-free survival; PPD, quality of life; PR, partial response; SAE, serious adverse event; TD, thalidomide and dexamethasone; TTP, time to progression; VCD, vincristine, cyclophosphamide, prednisone and doxorubicin; VCMCP/BADMCP/VBAD/B, vincristine, melphalan, cyclophosphamide, prednisone and doxorubicin; VCD, vincristine, doxorubicin, cyclophosphamide, prednisone; VD, vincristine, doxorubicin; VGP, very good partial response rate; VMP, bortezomib, melphalan and prednisone; VMPT, bortezomib, melphalan, prednisone and thalidomide; VTD, bortezomib, thalidomide and dexamethasone; VGPR, very good partial response rate; VMP, bortezomib, melphalan and prednisone; VMPT, bortezomib, melphalan, prednisone and thalidomide; VTD, bortezomib, thalidomide and dexamethasone.
| Medication                          | Trial name                  | Setting                              | Primary outcome | Setting | OS control | OS gain | OS HR | RR (DOR) | QOL | Toxicity | ESMO-MCBS score | ESMO-MCBS form | Reference(s) |
|------------------------------------|-----------------------------|--------------------------------------|----------------|---------|------------|---------|-------|----------|-----|----------|----------------|----------------|--------------|
| Dexamethasone ± lenalidomide       | CC-5013-MM-010              | Relapsed/refractory                  | TTP            | 4.7     | 6.6        | 0.35 (0.27– 0.48) | 20.6 | NA     | 0.66 (0.45– 0.96) |     | Improved (+1 point) | 3              | 2b            | 65           |
| Lenalidomide ± carfilzomib         | ASPIRE                      | Relapsed/refractory                  | PFS            | 17.6    | 8.7        | 0.69 (0.57– 0.84) | 40.4 | 7.9    | 0.79 (0.67– 0.96) |     | Slightly increased | 4              | 2a            | 66-67        |
| Lenalidomide ± ixazomib            | TOURMALINE-MM1              | Relapsed/refractory                  | PFS (interim)  | 14.7    | 5.9        | 0.74 (0.59– 0.94) |     | Immature | Not improved |     |           | 3              | 2b            | 68           |
| Lenalidomide ± daratumumab         | POLLUX                      | Relapsed/refractory                  | PFS (interim)  | 18.4    | 16.0       | 0.37 (0.27– 0.52) |     | Immature |           | Higher haematological toxicities | 3              | 2b            | 69           |
| Dexamethasone ± bortezomib         | APEX                        | Relapsed/refractory                  | TTP            | 3.5     | 2.7        | 0.55 (p=0.001) | 23.7 | 6.1    | 0.77 (p=0.027) |     |             | 3              | 2b            | 92-95        |
| Carfilzomib-d versus bortezomib-d  | ENDEAVOR                    | Relapsed/refractory                  | PFS            | 9.4     | 9.3        | 0.53 (0.44– 0.65) | 40   | 7.6    | 0.79 (0.65– 0.96) |     | Improved (abstract only) | Slightly higher SAEs | 3              | 2a            | 94-95        |
| Bortezomib ± daratumumab           | CASTOR                      | Relapsed/refractory                  | PFS            | 7.1     | 9.6        | 0.31 (0.24– 0.39) |     | Immature |           | Higher haematological toxicity | 3              | 2b            | 96-97        |
| Bortezomib ± spanobinostat         | PANORAMA1                   | Relapsed/refractory                  | PFS            | 8.1     | 3.9        | 0.63 (0.52– 0.76) | 30.4 | 3.25   | Immature | 3% increase in PN grade ≥3 (~1 point) | 2              | 2b            | 98           |
| Dexamethasone ± pomalidomide       | MM-003                      | Relapsed/refractory                  | PFS            | 1.9     | 2.1        | 0.48 (0.39– 0.60) | 8.1  | 4.6    | 0.74 (0.56– 0.97) |     |           | 4              | 2a            | 99           |
| Pomalidomide ± cyclophosphamide    | MMC-16705                   | Relapsed/refractory                  | ORR            | 4.4     | 5.1        | NS       |      |        | 64.7% vs 38.9%, gain 25.8% | 2              | 2c            | 100          |
| Daratumumab                        | SIRIUS                      | Relapsed/refractory                  | ORR            | 3.7     |            | 29% (7.4 months) |      |        |           |           | 3              | 3              | 101          |
| Daratumumab ± pomalidomide         | GEN501                      | Relapsed/refractory                  | Safety (16 mg/kg) | 5.6     |            | 36% (NR) |      |        |           |           | 3              | 3              | 102          |
| Daratumumab + pomalidomide ± d     | MMY1001                     | Relapsed/refractory                  | Safety         | 8.8     |            | 60% (>13 months) |      |        |           |           | 3              | 3              | 103          |
| Pomalidomide ± bortezomib + d      | MC1082                      | Relapsed/refractory                  | ORR            | 13.7    |            | 86%     |      |        |           |           | 3              | 3              | 104          |

<sup>d</sup> dexamethasone; DOR, duration of response; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; NA, not applicable; NR, not reached; NS, not significant; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; PN, polyneuropathy; QOL, quality of life; RR, response rate; SAEs, serious adverse events; TTP, time to progress.
Table 9  Feasibility testing of the ESMO-MCBS v1.1 for myelodysplastic syndrome (n=10)

| Medication                          | Trial name   | Setting                                      | Primary outcome | PFS control | PFS gain | PFS HR | OS control | OS gain | OS HR | RR (DOR) | QOL | ESMO-MCBS score | ESMO-MCBS form | Reference |
|-------------------------------------|--------------|----------------------------------------------|-----------------|-------------|----------|--------|------------|---------|-------|-----------|-----|----------------|----------------|-----------|
| Azacitidine versus SOC              | AZA-MDS-001  | High-risk MDS                                | OS              | 15 months   | 9.5 months | 0.58 (0.43–0.77) | 4       | 2a    | 105       |     |                |                |           |
| Decitabine versus SOC               | MDS FAB (IPSS ≥0.5) | Coprimary ORR and PFS                        | 7.8 months     | 4.3 months  | 0.58 (0.37–0.91) | Improved (+1 point) | 4       | 2b    | 106       |     |                |                |           |
| Lenalidomide (10mg/5mg) versus SOC  | LEN-MDS-004  | Transfusion-dependent patients with low-risk/intermediate-risk MDS del5q (IPSS ≥1) | RR (RBC-TI)    | 10 mg       |          |        |            |         |       |           |     |                |                |           |
| Lenalidomide versus SOC             | LEN-MDS-005  | MDS-WHO (IPSS <1)                            | RR (RBC-TI at ≥8 weeks) | 5 mg       |          |        |            |         |       |           |     |                |                |           |
| Antithymocyte globulin versus SOC   | SAKK 33/99   | MDS <10% bone marrow blasts                  | RR at 6 months |          |          |        |            |         |       |           |     |                |                |           |
| rHuEPO versus SOC                   | ICSG         | MDS <10% bone marrow blasts                  | RR (T)         |          |          |        |            |         |       |           |     |                |                |           |
| rHuEPO versus ±GCSF                 | GFM          | MDS <10% bone marrow blasts                  | RR (IWG 2006 modified) |          |          |        |            |         |       |           |     |                |                |           |
| EPO versus SOC                      | E1996        | MDS <10% bone marrow blasts                  | RR (WG 2000 modified) |          |          |        |            |         |       |           |     |                |                |           |
| rHuEPO +GCSF versus SOC             | GFM          | MDS <10% bone marrow blasts                  | RR (WG 2006 stringently modified) |          |          |        |            |         |       |           |     |                |                |           |
| Darbepoetin versus SOC              | MDS-WHO IPSS <1 | RBC transfusion incidence                    | 69% vs 36%, gain 23% |          |          |        |            |         |       |           |     |                |                |           |

del5q, S dilution; DOR, duration of response; ESMO-MCBS v1.1, European Society for Medical Oncology—Magnitude of Clinical Benefit Scale, version 1.1; FAB, French–American–British class; SOC, standard of care; IPSS, International Prognostic Scoring System; IWG, International Working Group; MDS, myelodysplastic syndrome; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RBC-TI, red blood cell transfusion independence; rHuEPO, recombinant human erythropoietin; RR, response rate; SOC, standard of care; TI, transfusion independence.
applicability of the ESMO-MCBS v1.1 as a first step to the development of a version validated for HMs.

There are several major differences in the behaviour of HMs as compared with solid tumour cancers. These differences arise largely from the more variable natural history of HMs which can range from fulminant (acute leukaemia and high-grade lymphomas) to almost benign (low-grade MDS). Furthermore, many of these malignant haematological diseases, even when they are not cured, they are characterised by very long PFS and OS that are rarely seen among incurable solid tumour malignancies. Finally, the endpoints used in the studies of treatments for HMs are sometimes different to those used in solid tumours and in some instances, such as CML, they are even disease-specific. Consequently, at the outset of this project we did not know if ESMO-MCBS v1.1 could be applied to studies in HMs, and if the grading of studies would generate grades considered reasonable by experts in the relevant diseases.

This evaluation of the behaviour of the ESMO-MCBS v1.1 in the grading of 80 studies across the full spectrum of HMs has demonstrated that the ESMO-MCBS v1.1 is widely applicable for the overwhelming majority of analysed studies (90% scoreable studies) and that the generated scores were generally adjudicated by clinical experts to reasonably accord with their evaluation of the magnitude of clinical benefit. In 5 of the 80 studies (6%), the ESMO-MCBS could not be applied at all and in 3 more studies (4%), it could not be applied to one of the evaluable parameters. In the evaluation of imatinib in CML, it generated scores that were considered to under-represent the true value of the intervention in the opinion of experts in the evaluated diseases.

Based on the analysis of the scorability of studies and the reasonableness of the generated results, this field testing identified six shortcomings in the current version of the ESMO-MCBS that will require redress to improve the applicability and reasonableness of ESMO-MCBS scoring for malignant haematological conditions.

1. Regarding single-arm studies with curative intent, such as CAR T-cell salvage therapies, the ESMO-MCBS v1.1 does not have a form to grade single-arm treatments with curative intent.
2. Regarding relatively indolent conditions with a very long PFS (or EFS) or OS such as CLL, CML, indolent lymphoma and MM, there is no mechanism to credit strong interim gains when the median of the control arm has not yet been reached.
3. The capitation of PFS at a maximal preliminary grade of 3, with provision for an upgrade based on tail of the curve only when there is a plateau in the arm with the study medication, may undervalue treatments with substantial late PFS gain but with no plateauing of the curves.
4. Regarding the standard molecular surrogate endpoints used for CML, the surrogacy of complete cytogenic response and level 4–5 MMR must be acknowledged and incorporated.
5. The scale does not make provision for the grading of non-inferiority studies based on response rate criteria.
6. In studies evaluating response rate as a primary end-point, there is no provision of QOL bonus if improved QOL is demonstrated as a secondary outcome.

Finally, it must be acknowledged that the results of the scale may not be reasonable for some of the least malignant of the HMs such as low-risk MDS. Most of the studies for MDS were evaluated based on response rates, but there was heterogeneity of the available definitions of haematological response and their clinical meaningfulness. This underlines the need for a stand-alone form regarding studies with such heterogeneity in their response rates.

ESMO and the EHA are committed to the development of a version of the ESMO-MCBS that is validated for HMs. Based on the findings of this study, a revised version of the ESMO-MCBS will be developed to address the identified shortcomings in the current version of the scale regarding the assessment of HMs. This development process will incorporate all the usual stringencies for accountability of reasonableness that have characterised the development of the ESMO-MCBS. This, thus far, included field testing, statistical modelling, evaluation for reasonableness and openness to appeal and revision. Applying such a scale will support future decision-making and will provide insights that could be helpful in the design of future clinical trials.

Author affiliations

1Department of Medicine I, Clinical Division of Oncology, Medical University of Vienna, Vienna, Austria
2Cancer Pain and Palliative Medicine Service, Department of Medical Oncology, Shaare Zedek Medical Center, Jerusalem, Israel
3Department of Hematology, Hospital Saint-Louis, Paris, Ile-de-France, France
4Adolescent and Young Adult Hematology Unit, Diderot University Paris Faculty of Medicine, Paris, Ile-de-France, France
5European Hematology Association, Den Haag, Zuid-Holland, The Netherlands
6Laboratory of Biostatistics, School of Health Sciences, National and Kapodistrian University of Athens, Athens, Greece
7Frontier Science Foundation-Hellas, Frontier Science Foundation-Hellas, Athens, Greece
8Department of Medical Oncology, University Medical Center Groningen, Groningen, The Netherlands
9Strategic Research Program on CLL, Division of Experimental Oncology, IRCCS Ospedale San Raffaele, Milano, Lombardia, Italy
10Universita Vita e Salute San Raffaele, Milano, Lombardia, Italy
11Department of Hematology/Oncology, Goethe University, Frankfurt am Main, Hessen, Germany
12Department of Hematology and Instituto de Investigacion Biomédica de Salamanca-IBSAL, University Hospital of Salamanca, Salamanca, Castilla y León, Spain
13Cambridge Stem Cell Institute, Department of Haematology, University of Cambridge, Cambridge, UK
14Department of Medicine I, Clinical Division of Hematology and Hemostaseology, Medical University of Vienna, Vienna, Austria
15Head Office, European Society for Medical Oncology, Lugano, Switzerland
16Department of Molecular Medicine, University of Pavia, Pavia, Lombardia, Italy
17Department of Hematology Oncology, Fondazione IRCCS Poliambulanza San Matteo, Pavia, Lombardia, Italy
18Department of Hematology, VU University Medical Centre Amsterdam, Amsterdam, Noord-Holland, The Netherlands
19Department of Hematology, Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland
Ibrutinib as initial therapy versus chlorambucil in older patients with treatment-naive CLL (RESONATE-2TM). Clinical Lymphoma Myeloma and Leukemia 2016:16.

Byrd JC, Brown JR, O’Brien S, et al. Ibrutinib versus ofatumumab in previously treated chronic lymphoid leukemia. N Engl J Med 2014;371:1101–11.

Brown JR, Hillmen P, O’Brien S, et al. Extended follow-up and impact of high-risk prognostic factors from the phase 3 RESONATE study in patients with previously treated CLL/SLL. Leukemia 2018;32:83–91.

Seymour JF, Ettinghausen SE, et al. Venetoclax–Rituximab in relapsed or refractory chronic lymphocytic leukemia. N Engl J Med 2018;378:1107–20.

Cortes JE, Saglio G, Kantarjian HM, et al. Venetoclax-Rituximab in relapsed or refractory chronic lymphocytic leukemia on imatinib versus interferon alfa plus low-dose cytarabine: results from the iris study. JCO 2016;34:235–42.

For patients with relapsed or refractory chronic-phase chronic myeloid leukemia. N Engl J Med 2003;348:994–1004.

Hahn EA, Glendenning GA, Sorensen MV, et al. Quality of life in patients with newly diagnosed chronic phase chronic myeloid leukemia on imatinib versus interferon alfa plus low-dose cytarabine: results from the iris study. JCO 2003;21:2138–46.

Hochhaus A, Larson RA, Guilhot F, et al. Long-term outcomes of imatinib treatment for chronic myeloid leukemia. N Engl J Med 2017;376:917–27.

Saglio G, Kim HW, Issajugiré S, et al.Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. N Engl J Med 2010;362:2251–9.

Hochhaus A, Saglio G, Hughes TP, et al. Long-term benefits and risks of front-line nilotinib versus imatinib for chronic myeloid leukemia in the chronic phase: 5-year update of the randomized ENESTnd trial. Leukemia 2016;30:1044–54.

Kantarjian H, Shah NP, Hochhaus A, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. N Engl J Med 2010;362:2260–70.

Cortes JE, Saglio G, Kantarjian HM, et al. Final 5-year study results of DASISION: the dasatinib versus imatinib study in treatment-naive chronic myeloid leukemia patients. JCO 2016;34:2333–40.

Cortes JE, Gambacorti-Passerini C, Deininger MW, et al. Bosutinib versus imatinib in newly diagnosed chronic myeloid leukemia: results from the randomized BFORE trial. JCO 2018;36:231–7.

Guilhot J, Preudhomme C, Mahon FX, et al. Analyzing molecular response in chronic myeloid leukemia clinical trials: pitfalls and golden rules. Cancer 2015;121:490–7.

Cross NCP, White HE, Müller MC, et al. Standardized definitions of molecular response in chronic myeloid leukemia. Leukemia 2012;26:2172–5.

Marcus R, Davies A, Ando K, et al. Obinutuzumab for the first-line treatment of follicular lymphoma. N Engl J Med 2017;377:1331–44.

Rummel MJ, Niederle N, Maschmeyer G, et al. Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. The Lancet 2013;381:1203–10.

Sehn LH, Chua N, Mayer J, et al. Obinutuzumab plus bendamustine versus bendamustine monotherapy in patients with rituximab-refractory indolent non-Hodgkin lymphoma (GADOLIN; a randomised, controlled, open-label, multicentre, phase 3 trial. Lancet Oncol 2016;17:1081–92. doi:10.1016/S1470-2045(16)30097-3.

Cheson BD, Trask DL, Cibinogen JC, et al. Health-Related quality of life and symptoms in patients with rituximab-refractory indolent non-Hodgkin lymphoma treated in the phase III GADOLIN study with obinutuzumab plus bendamustine versus bendamustine alone. Ann Hematol 2017;96:253–9.

Cheson BD, Chua N, Cervera J, et al. Overall survival benefit in patients with Rituximab-Refractory indolent non-Hodgkin lymphoma who received Obinutuzumab plus bendamustine induction and Obinutuzumab maintenance in the GADOLIN study. JCO 2018;36:2259–69.

Dreyling M, Jurczak W, Jerkeman M, et al. Ibrutinib versus temsirolimus in patients with relapsed or refractory mantle-cell lymphoma: an international, randomised, open-label, phase 3 study. The Lancet 2016;387:770–8.

Trníný M, Lampy T, Wenzl S, et al. Lenalidomide versus investigator’s choice in relapsed or refractory mantle cell lymphoma (MCL-002; sprint): a phase 2, randomised, multicentre trial. Lancet Oncol 2016;17:319–31.

Wang ML, Rule S, Martin P, et al. Targeting Btk with ibrutinib in relapsed or refractory mantle-cell lymphoma. N Engl J Med 2013;369:507–16.

Noy A, de Vos S, Thieblemont C, et al. Targeting Bruton tyrosine kinase with ibrutinib in relapsed/refractory marginal zone lymphoma. Blood 2017;129:2224–32.

Gopal AK, de Vos S, et al. PI3K5 inhibition by idelalisib in patients with relapsed indolent lymphoma. N Engl J Med 2014;370:1008–18.

Chen R, Zinzani PL, Fanale MA, et al. Phase II study of the efficacy and safety of pembrolizumab for relapsed/refractory classic non-Hodgkin lymphoma. JCO 2018;36:2259–66.

Younes A, Santoro A, Shipp M, et al. Nivolumab for classical Hodgkin’s lymphoma after failure of both autologous stem-cell transplantation and brentuximab vedotin: a multicentre, multicohort, single-arm phase 2 trial. Lancet Oncol 2016;17:1285–94.

Armand P, Engert A, Younes A, et al. Nivolumab for relapsed/ refractory classic Hodgkin lymphoma after failure of autologous hematopoietic cell transplantation: extended follow-up of the Multicohort single-arm phase II CheckMate 205 trial. JCO 2016;34:1428–39.

Pfreundschuh M, Trümper L, Österborg A, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial Alliance (MINT) group. Lancet Oncol 2011;12:1013–22.

Coiffier B, Thieblemont C, Van Den Neste E, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab–CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d’Etudes des Lymphomes de la Base Adoulinee. Blood 2002;101:1840–95.

Coiffier B, Lapage E, Brière J, et al. Chemoimmunotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N Engl J Med 2002;346:235–42.

Thieblemont C, Tilly H, Gomez da Silva M, et al. Lenalidomide maintenance compared with placebo in responding elderly patients with diffuse large B-cell lymphoma treated with first-line rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. JCO 2017;35:2473–81.

Crump M, Kuruvilla J, Cobian S, et al. Randomized comparison of gemcitabine, dexamethasone, and cisplatin versus dexamethasone, cytarabine, and cisplatin chemotherapy before autologous stem-cell transplantation for relapsed and refractory aggressive lymphomas: NCIC-CTG LY12. JCO 2014;32:3490–6.

Fettergell R, Coiffier B, Narayan G, et al. Pixantrone dimaleate versus other chemotherapy agents as a single-agent salvage treatment in patients with relapsed or refractory aggressive non-
Hodgkin lymphoma: a phase 3, multicentre, open-label, randomised trial. *Lancet Oncol* 2012;13:696–706.

70 Neelapu SS, Locke FL, Bartlett NL, et al. Axicabtagene Ciloleucel CAR T-cell therapy in refractory large B-cell lymphoma. *N Engl J Med* 2017;377:2553–64.

71 Schuster SJ, Bishop MR, Tam CS, et al. Tasigeneleucel in adult relapsed or refractory diffuse large B-cell lymphoma. *N Engl J Med* 2018.

72 Czuczman MS, Timney M, Davies A, et al. A phase 2/3 multicenter, randomized, open-label study to compare the efficacy and safety of lenalidomide versus investigator’s choice in patients with relapsed or refractory diffuse large B-cell lymphoma. *Clin Cancer Res* 2017;23:4127–37.

73 Assouline SE, Nielsen TH, Yu S, et al. Axicabtagene Ciloleucel for refractory large B-cell lymphoma: a phase 3 trial. *Lancet* 2019;393:2180–94.

74 Jacobsen ED, Sharma JP, Oki Y, et al. Brentuximab vedotin demonstrates objective responses in a phase 2 study of relapsed/refractory DLBCL with variable CD30 expression. *Blood* 2015;125:394–402.

75 Wilson WH, Young RM, Schmitz R, et al. Targeting B cell receptor signaling with ibritinib in diffuse large B cell lymphoma. *Nat Med* 2015;21:922–6.

76 Rosioli L, Oriol A, Teruel AI, et al. Superiority of bortezomib, thalidomide, and dexamethasone (VTD) as induction pretransplantation therapy in multiple myeloma: a randomized phase 3 PHEMA/GEM study. *Blood* 2012;105899–96.

77 Cavo M, Tacchetti P, Patriarca F, et al. Bortezomib with thalidomide plus dexamethasone compared with thalidomide plus dexamethasone as induction therapy before, and consolidation therapy after, double autologous stem-cell transplantation in newly diagnosed multiple myeloma: a randomised phase 3 study. *The Lancet* 2010;376:2075–85.

78 Mai EK, Bertsch U, Dürig J, et al. Phase III trial of bortezomib, cyclophosphamide, and dexamethasone (VCD) versus bortezomib, doxorubicin and dexamethasone (pad) in newly diagnosed multiple myeloma. *Leukemia* 2015;29:1721–9.

79 San Miguel JF, Schlag R, Khugaeva NK, et al. Persistent overall survival benefit and no increased risk of second malignancies with bortezomib-melphalan-prednison versus melphalan-prednison in patients with previously untreated multiple myeloma. *JCO* 2013;31:448–55.

80 San Miguel JF, Schlag R, Khugaeva NK, et al. Bortezomib plus melphalan and prednison for initial treatment of multiple myeloma. *N Engl J Med* 2011;365:906–17.

81 Palumbo A, Bringhen S, Rossi D, et al. Bortezomib-melphalan-prednisone-thalidomide followed by maintenance with bortezomib-thalidomide compared with bortezomib-melphalan-prednisone versus melphalan-prednison in patients with previously untreated multiple myeloma. *JCO* 2013:4144–55.

82 Benboubker L, Dimopoulos MA, Dispenzieri A, et al. Lenalidomide and dexamethasone in transplant-ineligible patients with myeloma. *N Engl J Med* 2014;371:906–17.

83 Mateos M-V, Dimopoulos MA, Cavo M, et al. Daratumumab plus bortezomib, cyclophosphamide, and dexamethasone (pand) for untreated myeloma. *N Engl J Med* 2018;378:518–28.

84 Durie BGM, Hoering A, Abidi MH, et al. Bortezomib with lenalidomide and dexamethasone versus lenalidomide and dexamethasone alone in patients with newly diagnosed myeloma without intent for immediate autologous stem-cell transplant (SWOG S0777): a randomised, open-label, phase 3 trial. *The Lancet* 2017;389:519–27.

85 Dimopoulos M, Spencer A, Attal M, et al. Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. *N Engl J Med* 2007;357:2123–32.

86 Stewart AK, Rajkumar SV, Dimopoulos MA, et al. Randomized multicenter phase 2 study of pomalidomide, cyclophosphamide, and dexamethasone in relapsed refractory multiple myeloma. *Blood* 2016;127:2561–8.

87 Anderloni A, Alini M, Iacoponi S, et al. Second generation colony-stimulating factor is better than erythropoietin alone to treat anemia in low-risk myelodysplastic syndromes: results from a randomized single-centre study. *Ann Hematol* 2006;85:174–80.
112 Greenberg PL, Sun Z, Miller KB, et al. Treatment of myelodysplastic syndrome patients with erythropoietin with or without granulocyte colony-stimulating factor: results of a prospective randomized phase 3 trial by the eastern cooperative Oncology Group (E1996). Blood 2009;114:2393–400.

113 Casadevall Net al. Health, economic, and quality-of-life effects of erythropoietin and granulocyte colony-stimulating factor for the treatment of myelodysplastic syndromes: a randomized, controlled trial. Blood 2004;104:321–7.

114 Platzbecker U, Symeonidis A, Oliva EN, et al. A phase 3 randomized placebo-controlled trial of darbepoetin alfa in patients with anemia and lower-risk myelodysplastic syndromes. Leukemia 2017;31:1944–50.

115 European Hematology Association. Available:https://ehaweb.org/