Correcting the Conclusion in a Study of Frameworks for Measurement of Absolute or Relative Clinical Survival Benefit

To the Editor Saluja et al1 evaluated the correlation of the European Society for Medical Oncology’s Magnitude of Clinical Benefit Scale (ESMO-MCBS) and the American Society of Clinical Oncology’s Value Framework (ASCO-VF) with absolute and relative survival measures, including the algorithmically calculated restricted mean survival time (RMST) derived from published Kaplan-Meier curves. Their conclusion regarding the shared inadequacy of ASCO-VF and ESMO-MCBS is based heavily on an unsubstantiated premise regarding RMST as a gold standard for evaluating survival benefit, and it is inconsistent with the results.

Comparison of an established scale to an efficacy measure that is not a gold standard cannot confirm or annul the scale’s value.2 Despite its merits, RMST is not a gold standard measure; it is not incorporated into the CONSORT statement, not routinely measured and reported, and its value as an efficacy measure tested in place of the hazard ratio (HR) remains under investigation.3 Furthermore, the calculation of RMST derived from Kaplan-Meier curves is prone to error, especially in view of the low-resolution images and the internal variability of the extraction software used. Indeed, the mean absolute relative difference between our RMST calculations (RUD) and the authors’ (RS) is not trivial (mean|RUD-RS|/RUD = 14%).

In contrast, ESMO-MCBS version 1.1 scores are calculated without ambiguity in a transparent and reproducible manner based on published and readily available data and the CONSORT requirements.4,5 Scores reflect both absolute and relative benefit, prognostic weighting and adjustments accounting for nonproportionate gains (particularly at the tail of the curve). The moderate correlation between a composite measure of one of its components (eg, ESMO-MCBS to median survival) reflects that the composite includes additional independent information, providing further credence to the ESMO-MCBS.

In addition, we note 3 further issues: ESMO-MCBS scoring was incorrect for 15 of the 79 studies, despite online availability of correct scores for 10 of 15 studies (https://www.esmo.org/Guidelines/ESMO-MCBS). Second, because HR is the only survival benefit component of ASCO-VF, the correlation of a measure with itself (HR to HR), with the reported correlation of 1, is a futile comparison. Finally, the conflations between the ESMO-MCBS and the ASCO-VF in the conclusion is not consistent with the presented data that highlighted the substantially better correlation to RMST of the ESMO-MCBS version 1.1 plus tail of the curve (moderate ρ = 0.67) in contrast to that of the ASCO-VF (weak ρ = 0.40).

We emphasize that the results generated by the ESMO-MCBS have been compared with the best surrogate of a “gold standard” measure, the experience and assessment of the experts in the field, providing validation of its reasonableness.4,6 Consequently, for all the stated reasons, we reject the authors’ conclusions regarding the validity of the ESMO-MCBS as a measure of clinical benefit.

Urania Dafni, ScD
Nathan I. Cherny, MBBS, FRACP, FRCP, LLDr
Elisabeth G. E. de Vries, MD, PhD

Author Affiliations: National and Kapodistrian University of Athens, Athens, Greece (Dafni); Shaare Zelek Medical Center, Jerusalem, Israel (Cherny); University Medical Center Groningen, University of Groningen, Groningen, the Netherlands (de Vries).

Published Online: October 17, 2019. doi:10.1001/jamaoncol.2019.4091

Conflict of Interest Disclosures: Drs Dafni and Cherny serve as members of the European Society for Medical Oncology’s Magnitude of Clinical Benefit Scale working group. Dr de Vries serves as the chair of the European Society for Medical Oncology’s Magnitude of Clinical Benefit Scale working group and reports institutional financial support for her advisory role from Daiichi Sankyo, Merck, National Surgical Adjuvant Breast and Bowel Project, Pfizer, Sanofi, Synthon and institutional financial support for clinical trials or contracted research from Amgen, AstraZeneca, Bayer, Chugai Pharma, CytoimmunX Therapeutics, GI Therapeutics, Genentech, Nordic Nanovector, Radius Health, Roche, Synthon outside the submitted work.


In Reply We thank Dafni et al for their comments. Our intent was not to invalidate but to empirically explore the measurement characteristics of the frameworks.1 We also did not state that restricted mean survival time (RMST) is the “gold standard” measure for absolute survival. Restricted mean survival time was included in our primary analysis owing to its advantage over median survival in capturing the entire survival distribution. Recognizing that RMST is not currently widely reported, we included an additional comparison with median survival that confirmed a moderate correlation with the European Society for Medical Oncology’s preliminary magnitude of clinical benefit grades (ESMO-PMCBGs). Acknowledging that RMST calculations by reviewers may be subject to interrater variability, we advocated for direct reporting of RMST...
by clinical trial investigators with access to individual patient data.1

Although an average difference of 14% between our RMST calculations and those of Dafni et al may not appear trivial at first glance, it should be noted that 14% of the mean of the primary analysis RMST differences represents less than 2 weeks. When discussing value in oncology, measurement differences of less than 2 weeks are not typically considered clinically meaningful.2

We did not cross-check our scoring with online scores because the article was submitted for publication prior to the posting of detailed scorecards that present ESMO-PMCBGs. Previous research has also questioned the reproducibility of ESMO scores because absolute concordance of ESMO-MCBS across research groups has been poor (44%).3 Nevertheless, when incorporating the posted scores into our primary analysis, we observed similar results, with ESMO-PMCBGs maintaining a low correlation with both RMST difference (ρ = 0.38) and hazard ratio (HR) (ρ = 0.42), and a moderate correlation with median survival (ρ = 0.65). Regarding the ASCO-VF comparison to HR, a perfect correlation was expected and our purpose was to clearly highlight that ASCO-VF is a purely relative measure despite the intention of the framework demanding an absolute measurement of survival.

Our conclusion is largely based on the primary analysis that does not credit tail-of-the-curve (TOC) gains, a significant source of nonconvergence between the ASCO and ESMO frameworks.4 The criteria to capture long-term survival (represented by TOC) used by both frameworks may be perceived as arbitrary and do not agree between frameworks (κ = 0.01), resulting in potentially inadequate evaluation of true long-term survival.5 This further highlights the benefit of incorporating a mean-based metric (RMST) into the frameworks to more accurately capture the entire survival distribution, including the TOC.

Currently, ESMO-PMCBGs combine 2 surrogates of relative and absolute survival, HR and median survival, respectively. Hazard ratio is a reasonable relative survival surrogate under the proportional hazard assumptions, which cannot always be assumed.6 Median survival reflects absolute benefit albeit at a single point of the survival curve. Restricted mean survival time provides a direct measure of absolute benefit that can circumvent some of these surrogate limitations. As our knowledge of value measurement increases, we may wish to be innovative in our approach and should question if these traditional measures are still optimal. The ESMO and ASCO, through their frameworks, are in an ideal position to adopt and advocate the use of novel metrics such as RMST to further improve measurement of value.

Ronak Saluja, BSc
Matthew Cheung, MD
Kelvin K. W. Chan, MD, MSc, PhD

Corresponding Author: Kelvin K. W. Chan, MD, MSc, PhD, Sunnybrook Odette Cancer Centre, 2075 Bayview Ave, T-Wing, T2-058, Toronto, ON M2N 3E6, Canada (kelvin.chan@sunnybrook.ca).

Published Online: October 17, 2019. doi:10.1001/jamaoncol.2019.4132

Conflict of Interest Disclosures: None reported.


Tumor Mutation Burden—From Doubts to Concerns

To the Editor: The role of tumor mutation burden (TMB) in predicting cancer immunotherapy response is predicated on the well-supported hypothesis that the immune system is able to recognize and respond to cancer-specific neoantigens. Even as a bulk and indirect neoantigen surrogate, TMB has been increasingly accepted as a biomarker of response. In their recent Viewpoint, Addeo et al1 compellingly described the adoption of TMB as overzealous “groupthink,” with real-world implications for patients. The implications are broad, with TMB influencing internationally recommended treatment guidelines, US Food and Drug Administration approval, and clinical practice. We also note an expanding cohort of clinical trials (eg, NCT03668119, NCT03178552, NCT03519412) that are actively using TMB status as an inclusion criterion.

With this context in mind, Addeo et al1 highlighted multiple key caveats and challenges associated with TMB as a metric. Specifically, they pointed toward a lack of standardization among TMB assays, which we have found to be further complicated by analytical differences (eg, among somatic variant identification pipelines), raising concerns of robustness in quantitation.2 Addeo et al also noted that TMB is difficult to practically assess in many cases, with a substantial observed rate of assay failure, and, on average, a considerable delay to results.

Although Addeo et al1 appropriately asserted the critical importance of overall survival (OS) as an end point, TMB itself should not be considered as a predictor of OS. Instead, the statistical associations between the 2 factors are likely explained indirectly via TMB’s correlation with immunotherapy response, a strong predictor of OS, and not by TMB itself.2 That is to say, TMB does not predict OS in an immuno-