ESMO Magnitude of Clinical Benefit Scale v1.1

Instructions

1. There are 5 forms:

   - Evaluation form 1: for new approaches to adjuvant therapy or new potentially curative therapies

     Hyper mature data from studies that were un-blinded after compelling early results with subsequent access to the superior arm are contaminated, subsequently late intention to treat (ITT) follow-up data are not evaluable.

   - Evaluation form 2a: for therapies that are not likely to be curative with primary endpoint of OS with separate sheets for:

     - IF median OS with the standard treatment is ≤12 months
     - IF median OS with the standard treatment >12 months, ≤24 months
     - IF median OS with the standard treatment >24 months

   - Evaluation form 2b: for therapies that are not likely to be curative with primary endpoint PFS with separate sheets for:

     - IF median PFS with standard treatment ≤6 months
     - IF median PFS with standard treatment >6 months

   - Evaluation form 2c: for therapies that are not likely to be curative with primary endpoint other than OS or PFS or equivalent (non-inferiority) studies.

   - Evaluation form 3: for single-arm studies in "orphan diseases" and for diseases with "high unmet need" when primary outcome is PFS or ORR.

2. The highest grade of the ESMO-MCBS is A in the curative setting and this is restricted to new curative treatments; for non-curative indications 5 is the highest possible grade, yet sufficient to trigger rapid consideration for reimbursement is B and 4.

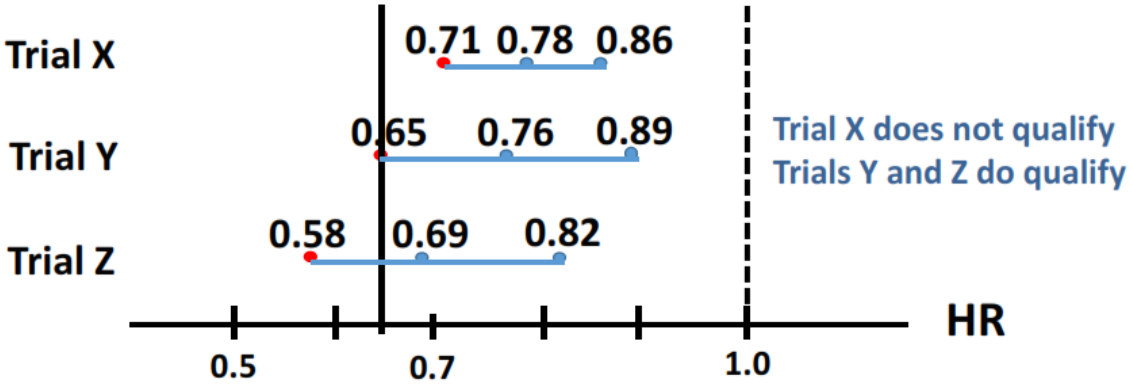
3. Analysis of phase III trials

   - Adequately powered studies showing statistically significant improvement in the primary outcome (defined by P <0.050)
   - Careful analyses "control arm" and identification of endpoints.
   - Check subgroup analysis.
     a. Studies with pre-planned subgroup analyses with a maximum of 3 subgroups can be graded (provided there is adjustment for multiple comparisons).
     b. When statistically significant results are reported for any subgroup, then each of these should be graded separately.
     c. Subgroups not showing statistically significant results are not graded.

d. Except for studies that incorporate collection of tissue samples to enable re-stratification based on new genetic or other biomarkers, findings from un-planned (post-hoc) subgroup analysis cannot be graded and they can only be used as foundation for hypothesis generation.

4. More than one outcome may be applicable
   - The statistical significance of secondary outcomes are determined by the same criteria as for primary outcomes i.e. (defined by *P* <0.05).

5. For a required HR, not the point estimate but the lower limit of 95% CI estimated based on the observed HR in the trial should encompass the required HR.



Example: for threshold set at HR ≤0.65 it is the lower limit of the 95%CI which has to be ≤0.65

6. In the Case of OS in the non-curative setting check for
   - Reduced toxicity
   - Improvement in quality of life
   - Report final adjusted grade taken into account toxicity, and QoL when relevant.

7. In case of PFS in the non-curative setting check for:
   - indicators of toxicity
   - survival data also available
   - early termination with crossover based on planned interim survival analysis
   - global QoL advantage using validated scale if applicable
   - report final adjusted grade taken into account toxicity, survival advantage and QoL when applicable.