Introduction

- Health technology assessment (HTA) bodies require lifetime estimates of mean costs and outcomes across relevant technologies.
- Cost-effectiveness analyses often use time-to-event data to estimate these costs and outcomes.
- Estimation of these data using parametric survival modelling is often necessary.
- Degree of influence of each treatment effect on cost effectiveness may be sensitive to the choice of scale for the parametric model.
- Parametric meta-analysis and network meta-analysis (NMA) are commonly used to estimate relative treatment effects for time-to-event endpoints, which should be based on systematic reviews of the clinical literature.
- These analyses often use published hazard ratios (HR) and the AFT assumption to estimate relative treatment effects.
- Time ratios (TRs), which are estimated from accelerated failure time (AFT) models, are recommended as an alternative measure of relative treatment effect when the PH assumption is violated, provided the associated AFT assumption holds.
- TRs are rarely reported in clinical trial publications.
- Methods now exist to reconstruct individual patient data (IPD) from published Kaplan-Meier (KM) plots, enabling assumptions around the scale of the treatment effect to be formally assessed and evidence such as TRs be used to estimate final model parameters.
- We utilise a case study in non-small cell lung cancer (NSCLC) to demonstrate the impact of choice of scale in NMA of time-to-event endpoints, in terms of mean survival.

Objectives

- To compare relative treatment effects and mean survival estimates obtained from a NMA in NSCLC, conducted on the PH and AFT scales.

Methods

- KM curves for survival (OS), derived from a subset of an evidence network analysis in a previous NICE appraisal (TA101), were digitised. The IPD for each study were then reconstructed from the KM plots.
- The subset of evidence was formed from published studies reporting an OS KM graph that connected together through common comparators.
- The PH assumption was assessed using log-cumulative hazard ratios, Schoenfeld residual plots, Schoenfeld residual global test; the AFT assumption was assessed using C-G plots.
- HRs and TRs were estimated for each study using a PH and AFT Weibull model and synthesised in a Bayesian NMA.
- Although not commonly reported in clinical trial publications, TRs can be estimated from KM plots or summary survival data, and a NMA synthesising TRs can be conducted using the log-cumulative hazard plot.
- The differences in mean survival estimates using the PH and AFT NMA results would lead to violations in both assumptions.

Results

- The KM plots for all studies based on the re-constructed IPD are presented in Figure 1. The OS data for this treatment is limited and it is difficult to determine which scale is more appropriate.
- The KM plots for all studies based on the re-constructed IPD are presented in Figure 1. The OS data for this treatment is limited and it is difficult to determine which scale is more appropriate.

- The use of AFT models does not solve the issue of non-constant treatment effects which would lead to violations in both assumptions.
- TRs do not solve the issue of non-constant treatment effects which would lead to violations in both assumptions.

Discussion

- The NMA was conducted on both the PH and AFT scales with only input data as HRs or TRs differing between analyses.
- The NMA was conducted on both the PH and AFT scales with only input data as HRs or TRs differing between analyses.
- The use of AFT models does not solve the issue of non-constant treatment effects which would lead to violations in both assumptions.
- TRs do not solve the issue of non-constant treatment effects which would lead to violations in both assumptions.

Limitations

- There were potential violations in both the PH and AFT assumptions across studies and therefore the formal assessment of assumptions alone did not provide clear evidence to justify the use of one scale over another.
- There were potential violations in both the PH and AFT assumptions across studies and therefore the formal assessment of assumptions alone did not provide clear evidence to justify the use of one scale over another.

Conclusions

- TRs are an alternative measure of relative treatment effect that can be considered in cases where the PH assumption does not hold, provided the AFT assumption is satisfied.
- TRs are an alternative measure of relative treatment effect that can be considered in cases where the PH assumption does not hold, provided the AFT assumption is satisfied.
- The use of AFT models does not solve the issue of non-constant treatment effects which would lead to violations in both assumptions.
- TRs do not solve the issue of non-constant treatment effects which would lead to violations in both assumptions.

References